

DISCLAIMER NOTICE



THIS DOCUMENT IS BEST
QUALITY AVAILABLE. THE COPY
FURNISHED TO DTIC CONTAINED
A SIGNIFICANT NUMBER OF
PAGES WHICH DO NOT
REPRODUCE LEGIBLY.

FINAL REPORT

Clarkson University, URI, Subcontract to BU

Prof. E. Barouch, Principle Investigator

Submitted to AFOSR.

Dr. A. Nachman, Program Manager

and

Clarkson University : Division of Research

Final Report

This is the final report for the four year URI to Clarkson University and a subcontract to Boston University.

During the duration of this project, we have developed several new algorithms for simulation of microstructures, devices, and fabrication methods for IC technology. We have interacted extensively with both industrial and government labs and collaborated with personnel of both.

We have developed algorithms for imaging of microstructures, investigated the effects of off-axis illumination and finite-thickness effects on mask imaging. Furthermore, we have investigated the effects of post-exposure baking on chemically amplified resists, optimized stepper parameters, as well as proximity effects on mask making. Last, but not least, we developed new visualization algorithms that allow multiple viewing of the microchip processes.

We have collaborated extensively with Air Force/Lincoln Labs personnel and have published three joint papers with them. We have also transferred our developments to the microchip facility of the NSA, Wright Patterson AFB, and the NRL. Our industrial collaboration with AMD, IBM, HP, DEC, LSI, Micon, and TI has continued and expanded.

The material in this report is organized as follows. One, a list of publications supported by this project. Two, a report on our new visualization algorithm; three, several selected papers supported under this project.

During the tenure of this project the ideas of the nonuniform grid FFT have been extended and the algorithm developed for its implementation integrated into our various simulation algorithms, including our Helmholtz solver. Furthermore, the exposure-kinetics of photoresist exposure with partially coherent light has been addressed and implemented, yielding simulations that are very close to the observed results.

We summarize this introduction with the comment that much progress has been made in the simulation process of microchip fabrication with the support provided from this grant.

Publications supported by the URI grant

- D. Cole, E. Barouch, U. Hollerbach, S. A. Orszag, "Extending scalar aerial image calculations to higher numerical apertures", *J. Vac. Sci. Technol.*, **B10**, 3037, 1992.
- D. Cole, E. Barouch, U. Hollerbach, S. A. Orszag, "Derivation and Simulation of Higher Numerical Aperture Scalar Aerial Images", *Jpn. J. Appl. Phys.* **31**, 4110, 1992.
- N. Eib, E. Barouch, U. Hollerbach, S. A. Orszag, "Characterization and Simulation of Acid Catalyzed DUV Positive Photoresist", *SPIE* **1926**, 1993.
- E. Barouch, D. Cole, U. Hollerbach, S. A. Orszag, "Vector Aerial Image with Off-Axis Illumination", *SPIE* **1927** 1993.
- E. Barouch, R. Coifman, C. Ma, M. Peckerar, V. Rokhlin, "Advanced Transformational Analysis Applied to E-Beam Proximity Effect Correction", *SPIE* **1924** 1993.
- W. A. Arnold, E. Barouch, U. Hollerbach, S. A. Orszag, "A focus vernier for optical lithography", *SPIE* **1925** 1993.
- M. Templeton, E. Barouch, U. Hollerbach, S. A. Orszag, "Suitability of high numerical aperture i-line steppers with oblique illumination for linewidth control in $0.35\ \mu\text{m}$ complex circuit patterns", *SPIE* **1927** 1993.
- R. R. Kunz, M. A. Hartney, R. W. Otten, E. Barouch, U. Hollerbach, *SPIE* **1927** 1993.
- Phase-Shift Mask Defect Printability Analyses Using 3D Aerial Imaging. BACUS, 1994
- Linewidth Control and Proximity Effects in I-line Steppers with Oblique Illumination, *Jpn. J. of Appl. Phys.*, 1995.
- E. Barouch, U. Hollerbach, and S. A. Orszag, "Defocus Asymmetry in Projection Printing", *J. Vac. Sci. Technol.* (appeared in Nov-Dec 1994 issue).
- E. Barouch, U. Hollerbach, and S. A. Orszag, "Aerial Image of 3D Phase-Shifted Recticles—3D FAIM", *Proceedings of SPIE* **2197** 442(1994).
- C. A. Spence, J. I. Nistler, E. Barouch, U. Hollerbach, and S. A. Orszag, "Automated Determination of CAD Layout Failures Through Focus: Experiment and Simulation", *Proceedings of SPIE* **2197** 302 (1994).
- C. A. Spence, J. I. Nistler, R. Vallishayee, E. Barouch, and U. Hollerbach, "Optimization of PSM Designs Including Defocus Effects", presented at *Proceedings of SPIE* (1995).
- E. Barouch, U. Hollerbach, and R. Vallishayee, "OPTIMASK: An OPC Algorithm for Chrome and Phase-Shift Mask Design", presented at *Proceedings of SPIE* (1995).
- U. Hollerbach, "Extraction of Zernike Coefficients from Measured Interferometric Wavefront Data", (in preparation).
- U. Hollerbach, "Prediction of Surface Swelling Profiles During Silylation", (in preparation).
- U. Hollerbach, "Comparative Effectiveness of Top and Bottom Anti-Reflective Coatings for Microlithography", (in preparation).
- U. Hollerbach, "Effects of Aliasing Errors on Microlithographic Image Computations", (submitted).

n2dX - A Multiple-Image Color Contour Plotter

"n2dX" is a color-contour plotting program designed to graphically display the results when several computational runs are made varying one parameter (defocus, for example) over a range of values. The output files for the run come out with extensions numbered sequentially starting at zero. In n2dX's main window, merely enter the "base name" (omitting the numerical extension), and the number of plots in the run.

"n2dX" will then load them all simultaneously. Its general operation is then the same as "2dX", except that zooming and slice plotting affect all plots simultaneously. On each plot's control panel there is a button that, when clicked, cycles among the images, essentially "animating" visually the effect of varying the parameter from one end of its range to the other.

There is also a button, labeled "MultiPlot", that will display that window's images side-by-side in a large window. This display can be terminated either by typing "q", or by closing the window,

A single plot, one not part of a series, can be viewed by entering its full name and giving "1" as the number of images.

Slices:

Another feature of "n2dX" is plotting the value of the function along a line segment, either horizontal or vertical. With the cursor, mark a rectangle having as one of its sides the line along which you want function values plotted. This line will then be either the "N" (North, or top), "E" (East, or right), "S" (South, or bottom), or "W" (West, or right) edge of this rectangle. Make sure that this rectangle is highlighted (Click on "Choose Region" until it is), click on the panel "Direction" button until the correct edge is selected. Click on "Plot Slice" and the function plot will appear next to the color contour plot.

While the cursor is in the slice-plot window, the x, y, and function value of that point will be displayed on the window's control panel, so values don't need to be read by eye off of the plot.

Grids:

One the main window, there are buttons selecting the gridding style. "No Grid", "Element Grid", or "Full Mesh" can be selected. The next time a window is redrawn, the appropriate grid lines will be displayed. It might be necessary to click in a window once or twice to force it to be redrawn.

OPC Corrections Algorithms with Resist Processing

Eytan Barouch and Uwe Hollerbach
Department of Manufacturing Engineering
Boston University
Boston, Mass.

Abstract

We have developed a new optical proximity correction (OPC) algorithm based in part on aerial image and in part on resist development. Resist development considerations include full 3D simulations, SEM information, as well as quick simple modeling. A new regula-falsi-based optimization scheme with multiple constraints has been introduced.

1. Introduction

In the past two years optical proximity corrections have been examined in great detail and have been reported in the literature (SPIE 1995, PMJ 1995). It was clear from the beginning that careless corrections, including line shortenings (LS), line narrowing (LN), and corner rounding (CR), can cause features to bridge, at least optically. To further complicate matters, a mask layout that appears nearly perfect at zero defocus can have very poor performance at end-range defocus, and vice-versa. Furthermore, depending on resist and layout, the aerial image serves at least as a very good indicator of how developed features will appear. The indicator is not uniform and in some cases processing (exposure, post-exposure bake, and dissolution) must be taken into account. Due to the vast variety of features and patterns, it is impossible to predict globally on a feature-by-feature basis the precision of the aerial image approximation to the processed resist. The typical range of influence is 2 to 4 microns, while feature sizes are below 0.5 microns, or roughly one wavelength in size.

In a previous set of publications (SPIE 95, PMJ 95) we have reported a new algorithm based entirely on a threshold aerial image model that in most cases works extremely well and that is based on a constrained Newton optimization scheme. However the algorithm assumes a priori an uncorrected mask. Furthermore, the aerial image (FAIM) on which it is based and the optical design rule checker (ODRC) have been parallelized. In certain tightly-packed lines (inverted transistor-like design) line shortening takes place, and if the line shortening is allowed to fully correct the aerial image bridges with the transverse line. Thus, an algorithm based on aerial image alone will require a redesign of the device's

assembly, which is an extremely undesirable situation, and which should be used only as an absolute last resort.

In situations of this sort, a much better guide to the OPC is undoubtedly the developed resist pattern, either simulated or SEMs. This is indeed the main criterion. We have therefore redeveloped major parts of the algorithm OPTIMASK and named it EOPTIMASK in honor of Eric Jackson, who played a significant role in developing this code.

EOPTIMASK has the capability of incorporating aerial image and a Gaussian convolution (or any other convolution) of the aerial image to represent a resist image, and the ability to take into account full 3D simulation as well as developed micrograph surfaces. It has been fully parallelized on Cray T3D and IBM SP2 platforms. In all cases it is completely linear, allowing full chip pre- and post-processing, using domain decomposition.

Several examples are given.

2. Description

EOPTIMASK accepts a FAIM input file containing fractured and flattened patterns of the CAD layout along with the optical parameters of the system. Furthermore, it optionally takes either a 3D developed resist surface or a convolution function that is convolved with the aerial image to approximately represent the developed surface.

Once the surface criterion and its level (threshold) is established the design rule checker is engaged to determine where the threshold exceeds CD and where the surface gradient is below acceptable contrast level. This takes place at both 0 defocus as well as at end-range defocus. The length and direction of each failure is recorded in three different types: line narrowing, line shortening, and corner rounding.

The line-narrowing segments are averaged over the defocus range. This is particularly critical in isolated lines where the aerial image has a tendency to spill out of the feature while the end-range defocus contour is embedded inside the feature. When that happens the mask cannot be corrected for this defocus range. As discussed in the introduction, the fashion in which line shortening takes place is strongly dependent on the physical interaction amongst the features. The diffraction patterns obtained from isolated and packed lines are very different. Therefore a complete global uniform line shortening correction algorithm is beyond our reach. Once the lengthening of the shortened line is achieved, the corner corrections are done via serifs whose lengths are determined by the failure and whose widths are dependent initially on the failure at the midpoint. Once the initial line movements have been determined, each movement is bracketed on both sides to determine optimal design under the restriction that two features cannot be closer to each other than a given value. A second set of restrictions prohibits correction on segments smaller than a multiple of the minimum feature size and restricts it to a fixed spot size, assuring that

the affected mask is manufacturable. The iterations take place, in principle, on all corrections. A regula-falsi type algorithm is imposed and it exits when the mask can't be further corrected. Finally, the algorithm selects the best of the last three masks and reports it.

3. Physical Considerations

The process of resist development involves three basic physical steps:

a) Exposure: This step is largely dominated by a Maxwell-Material equation coupled to the kinetics of the photoactive compound changes during exposure.

b) Post-Exposure Baking (PEB): This is largely controlled by diffusion or reaction-diffusion systems.

c) Dissolution (Resist Development): The physics of this process can be understood as a reaction-diffusion equation of the dissolver penetrating the resist film and removing the molecules that can be removed. From energy considerations of front tracking, a Hamilton-Jacobi principle is fully capable of describing the dissolution.

d) Plasma etching: This process can be described as a system of hyperbolic conservation laws.

In principle these basic processes transform the aerial image surface into a developed (or etched) surface pattern. Naively, one might consider convolving the aerial image with an integral operator to obtain the final resist surfaces. The speed of convolution and its simplicity make this operation very tempting, and in many cases it yields reasonable-looking results. However, the inherent problem remains: How does one know whether the particular mask design is truly represented by a convolution of the aerial image with a filter-like function? Furthermore, from a strict mathematical viewpoint the existence of such a convolution implies linear operation. If the process is truly linear, its ingredients (a, b, c, d above) must be linear as well. One might argue that b. is linear or nearly-linear, and a. is approximately linear, but neither the Hamilton-Jacobi energy balance nor the hyperbolic conservation law can be approximated by a linear system. In particular, the dissolution, in its simplest form, can be represented locally by the eikonal equation, where the square of the surface gradient equals the square of the dissolution front velocity. The same notions can easily be demonstrated for plasma etching systems. The representation of a highly-nonlinear system by a linear operator which works for some cases is a very dangerous one since it can lead to uncontrolled errors no matter how expertly chosen are the parameters of the convolution. However, at the small window level it may work, and a highly-windowed algorithm such as EOPTIMASK could serve as a good candidate to optimize an entire chip.

Clearly the best way to bias the mask is to have a large number of micrographs where correlations between mask design and printed features can be obtained by multiple-regression analysis. However, this is a very expensive and impractical way to proceed, and it can work at best at the local level. The same is true for 3D resist simulation since all computations in 3D are somewhat cumbersome and time-consuming. Therefore a local correlation can be tested for each window where the resist patterns are simulated locally. We believe that a global correlation function between aerial image and developed resist is unobtainable. Furthermore, attempting to find a linear operator that transforms the aerial image to the developed surface is an extremely ill-posed problem for which there is no guaranteed solution, and when a solution exists, it is almost certainly not unique. In other words, it is a procedure doomed to failure.

4. Examples

We concentrate in this section on demonstrating the concept of our discussion. We have chosen several test masks to illustrate the proximity effects and the connection of proximity corrections to 3D resist images. In Figure 1A we show the corrected mask of an inverted-like transistor system. The corrected mask has 0.4 micron lines, a space between them of .4 microns, and .5 micron distance to the outer edges. The proximity correction, once chosen as nonaggressive, corrects for all the line shortening. Figure 1B shows both the uncorrected and the corrected images superimposed on the layers. Figure 1C shows the corrected image alone, superimposed on the original design. As can be seen, a small amplitude oscillation begins to develop as the vertical lines attempt to approach the horizontal bar. This takes place at .6 microns defocus, rendering impossible the complete correction of this structure via aerial image. We have performed the full 3D simulation of this design, with the aerial image displayed in Figure 2A. In Figure 2B we show the developed resist images of the corrected image and it can be seen very clearly that a fair bit of the line shortening and the line narrowing displayed in Figure 2C at 0 defocus and Figure 2D at .6 microns defocus has disappeared. This resist image was obtained for the AZ-Spectralith resist, which was chosen for the sake of the current simulation for its extreme properties. Specifically, this resist happens to have a dissolution rate function with a sharp cutoff as measured by N. Eib and reported earlier.

In the next example, we consider a system of six parallel lines and spaces with .7 microns pitch. Using the Spectralith resist, we see the corner rounding at both 0 and end-range defocus. However, the line shortening and .6 microns defocus as well as the line narrowing are more severe, as expected. In Figures 3C and 3D we show the actual line shortening for a single isolated line at .4 microns feature size and 0 and .6 microns defocus. The conclusions that we draw from these images are very simple and very consistent. Isolated lines have the aerial image extended outwards at 0 defocus, while packed lines perform better. The situation at .6 microns defocus is reversed. The CD control is relatively good for isolated lines while it is beginning to degrade at end-range defocus. These figures illustrate unequivocally that even though automated proximity effects can improve the resist images, it is not a quick fix solution. The physics of the process must

be better understood to achieve better simulation. The images discussed so far come from very simple designs that contain most of the inherent difficulties associated with a large mask OPC.

Having both the aerial image and the resist surface, our attempt to find a precise linear integral operator was a failure. Also, the convolution with Gaussian variances of .01 resulted in significantly degraded resist images as compared with the actual simulations.

For consistency with our previous OPTIMASK code, we illustrate the results of EOPTIMASK on both the "nikon mask" and our standard SRAM test case. In both cases the result of EOPTIMASK are superior to the previous algorithm. These two images are displayed on Figures 4A and 4B.

Last but not least, in figure 5A we show an IBM test case of 62 by 85 microns with .35 feature size. In Figure 5B the corrected mask is illustrated. In Figure 5C we show the superimposed aerial image from the corrected mask where lim-opt played a major role. The failures are relatively high at .6 microns defocus and EOPTIMASK reduced the total failure area by about 95% in most of the cases. The relaxation of the lim-opt shows in Figure 5D clear bridging of the critical features. This bridging again cannot be corrected by a simple linear operation.

5. Conclusion

Optical Proximity Correction is a powerful tool to bias masks within the manufacturing range. However it is not always possible to achieve complete correction. The complexity of the design certainly affects the globality of the process and it is somewhat wishful thinking to correct a whole mask in one pass. In many cases EOPTIMASK works very well. In some cases it is insufficient and in some cases the mask needs to be redesigned. The user and the designer need to evaluate the results in order for successful layout to be manufactured.

6. Acknowledgements

The authors thank E. Jackson, S. Knodle, and S. Orszag for many fruitful discussions. The support of AFOSR, HPCD and ARPA is gratefully acknowledged.

7. References

- 1) D. Cole, E. Barouch, U. Hollerbach, and S. A. Orszag, "Calculation of Aerial Images for High Numerical Aperture Projection Systems", J. Vac. Sci. Technol. bf B10, 3037 (1992).

- 2) D. Cole, E. Barouch, U. Hollerbach, and S. A. Orszag, "Derivation and Simulation of Higher Numerical Aperture Scalar Aerial Images", Jap. J. Appl. Phys. bf 31 (48) 4110 (1992).
- 3) C. A. Spence, J. I. Nistler, E. Barouch, U. Hollerbach, and S. A. Orszag, "Automated Determination of CAD Layout Failures Through Focus: Experiment and Simulation", Proceedings of SPIE **2197** 302 (1994).
- 4) C. A. Spence, J. I. Nistler, R. Vallishayee, E. Barouch, and U. Hollerbach, "Optimization of PSM Designs Including Defocus Effects", presented at Proceedings of SPIE (1995).

image3.faim

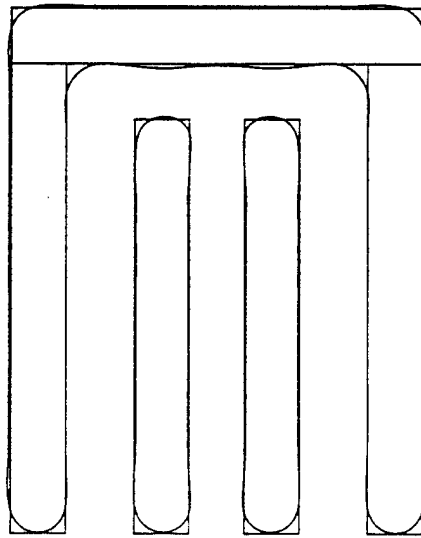


image3.faim

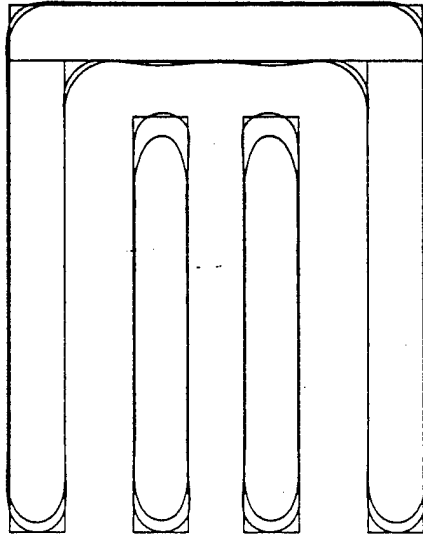
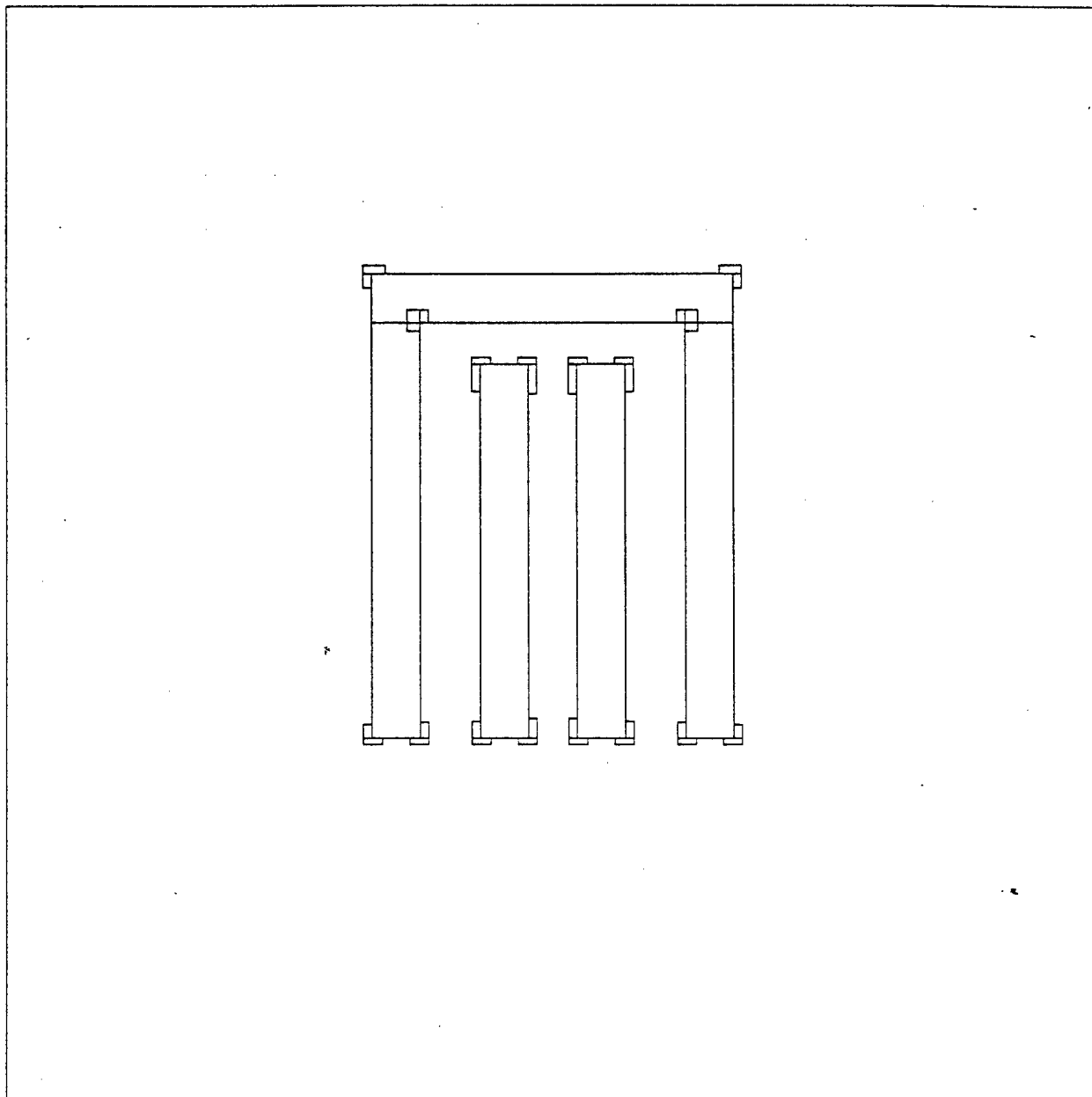
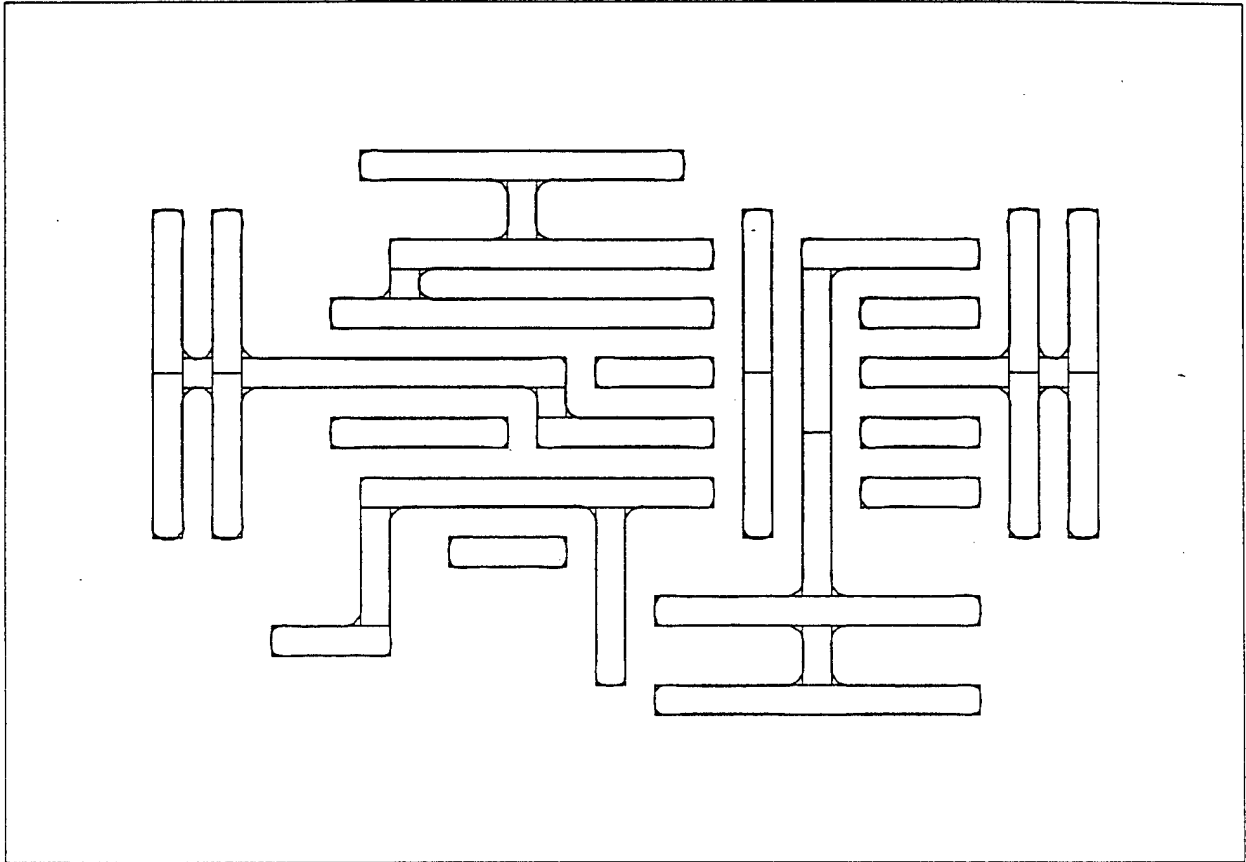


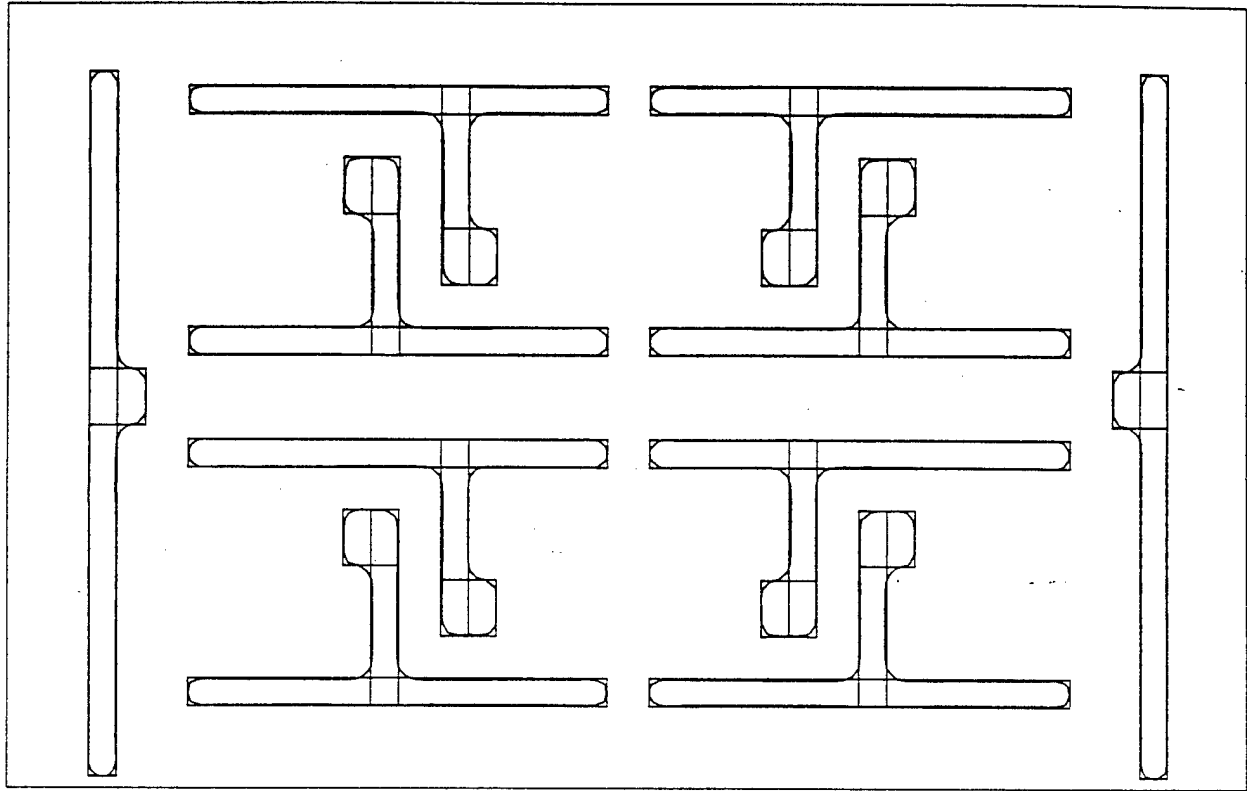
image3.optimask

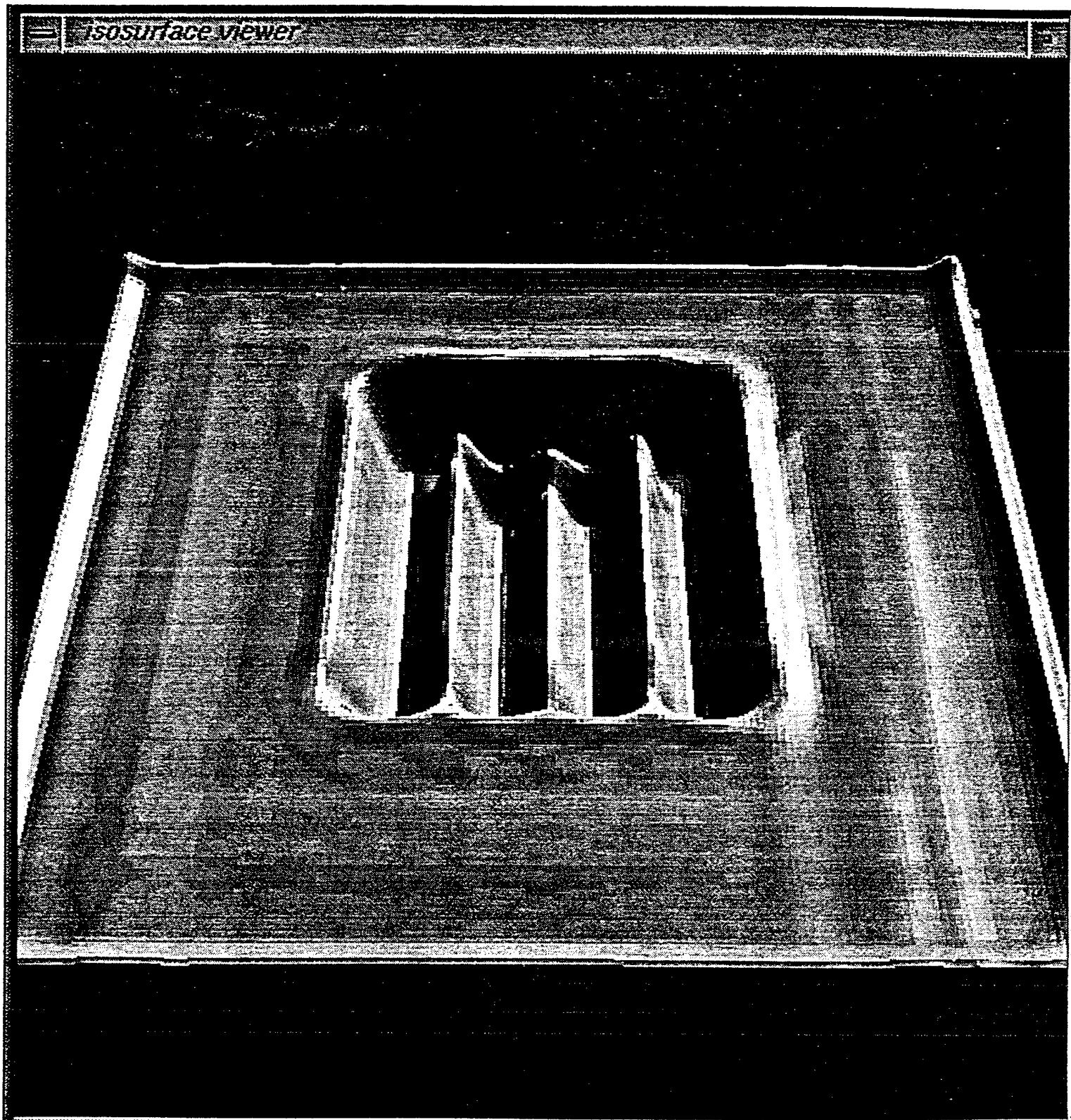


nikon.faim.1.4

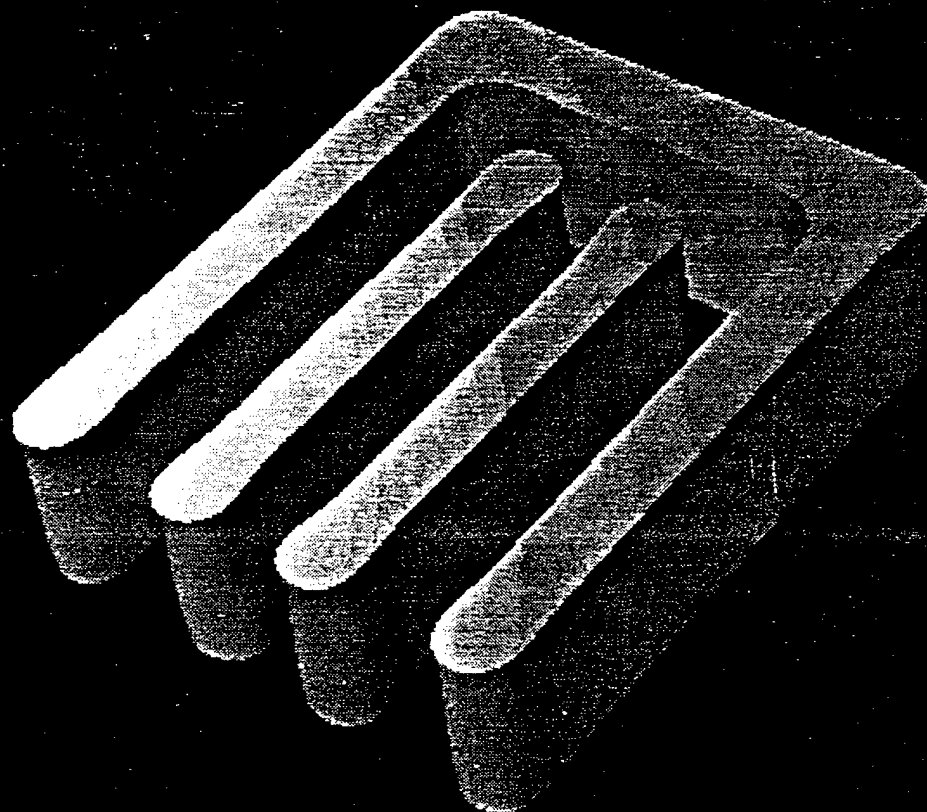


sram.iline.faim

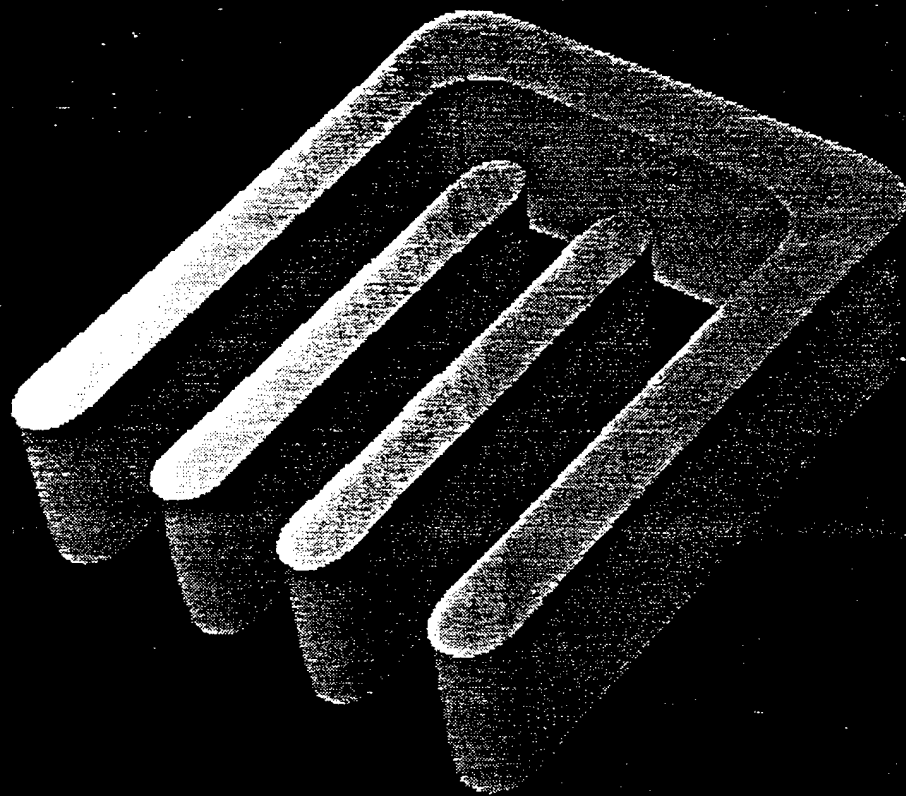




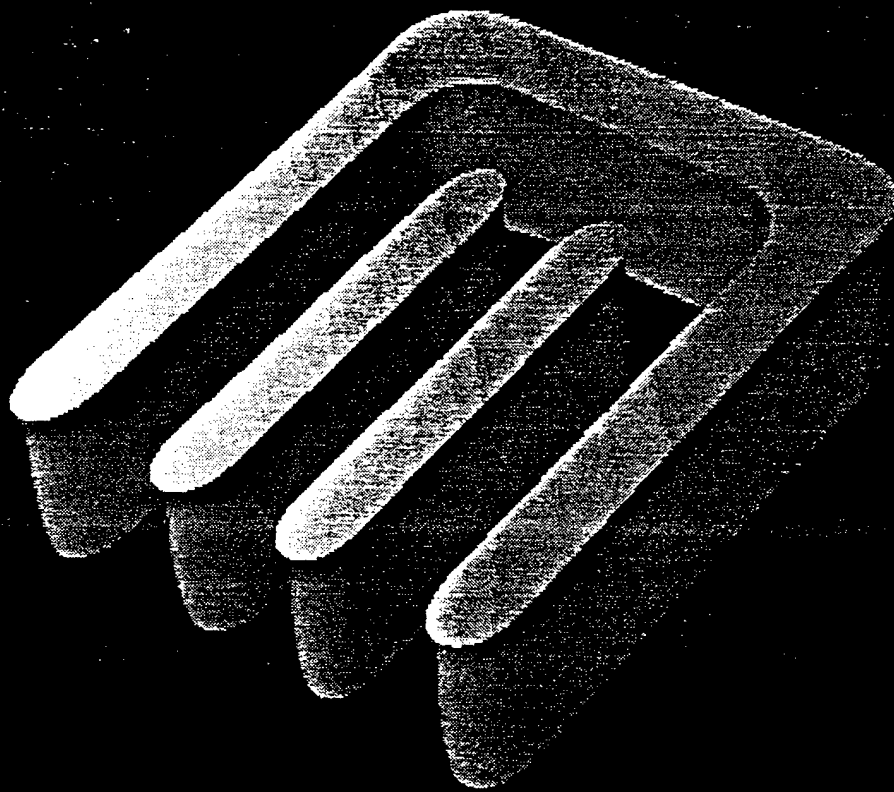
Inverted Transistor Line Short. Correction



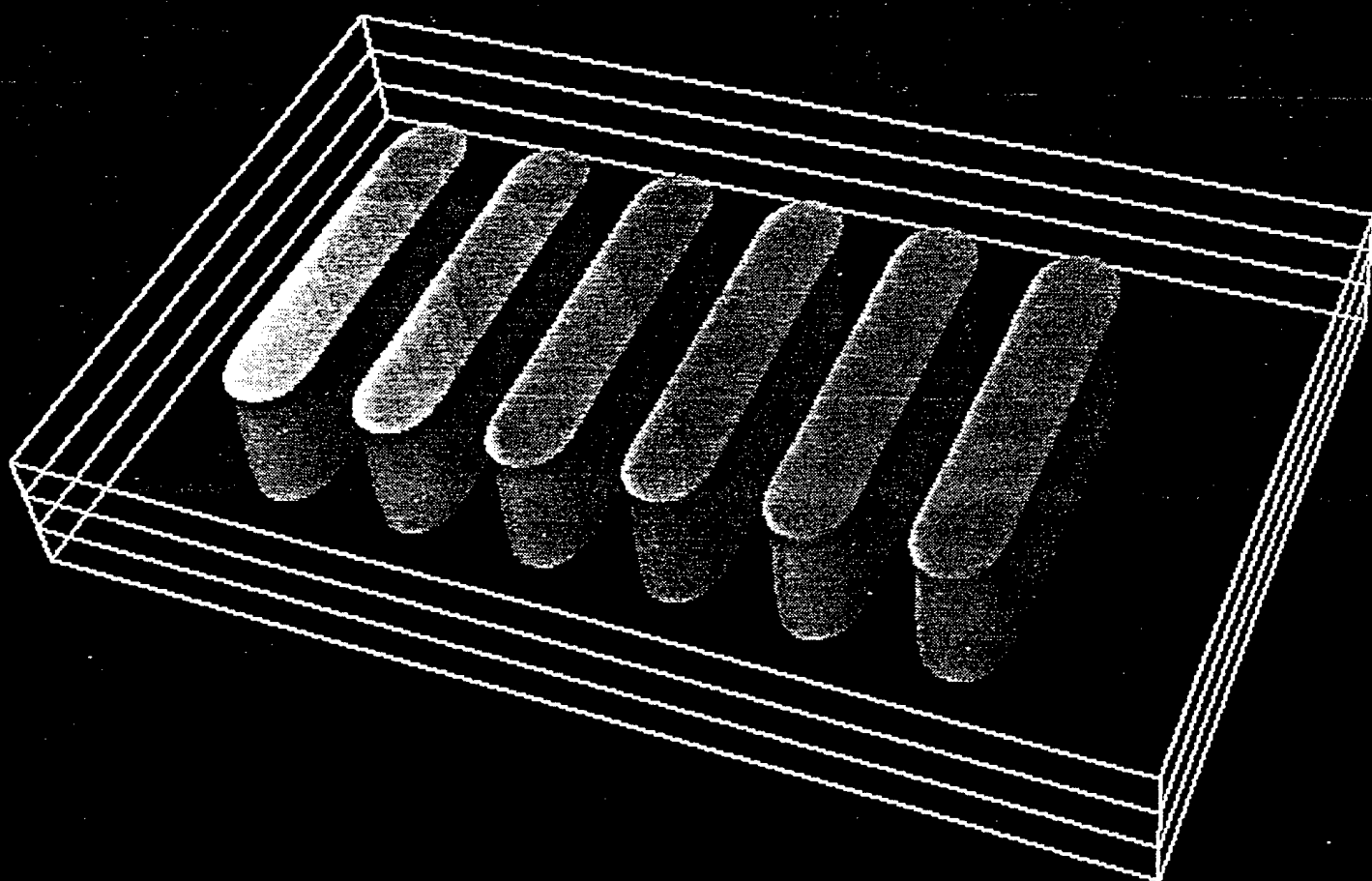
Inverted Transistor Line Shortening Odef



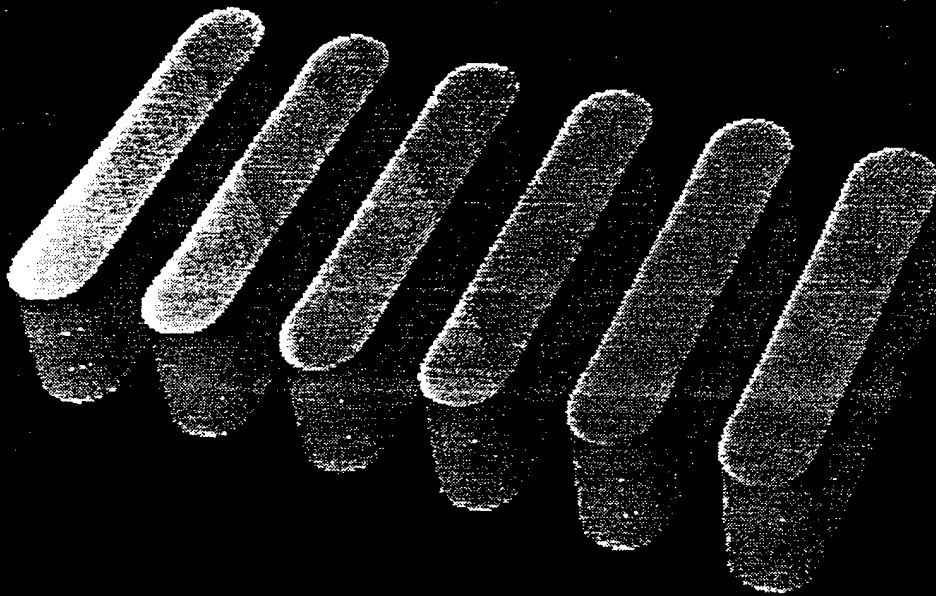
Inverted Transistor Line Shortening 0.6def



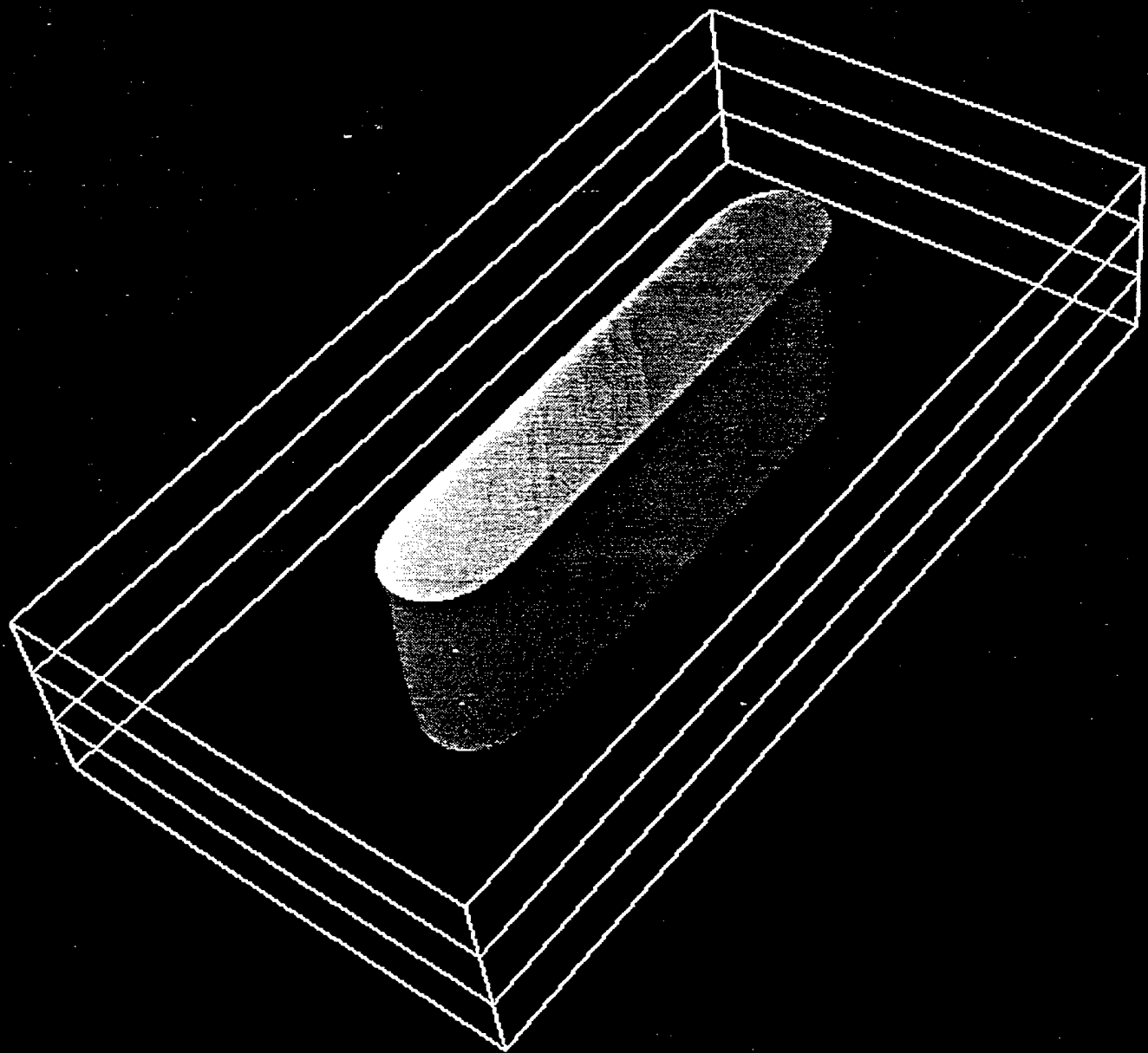
Shortening of L/S of .7u pitch0.6def



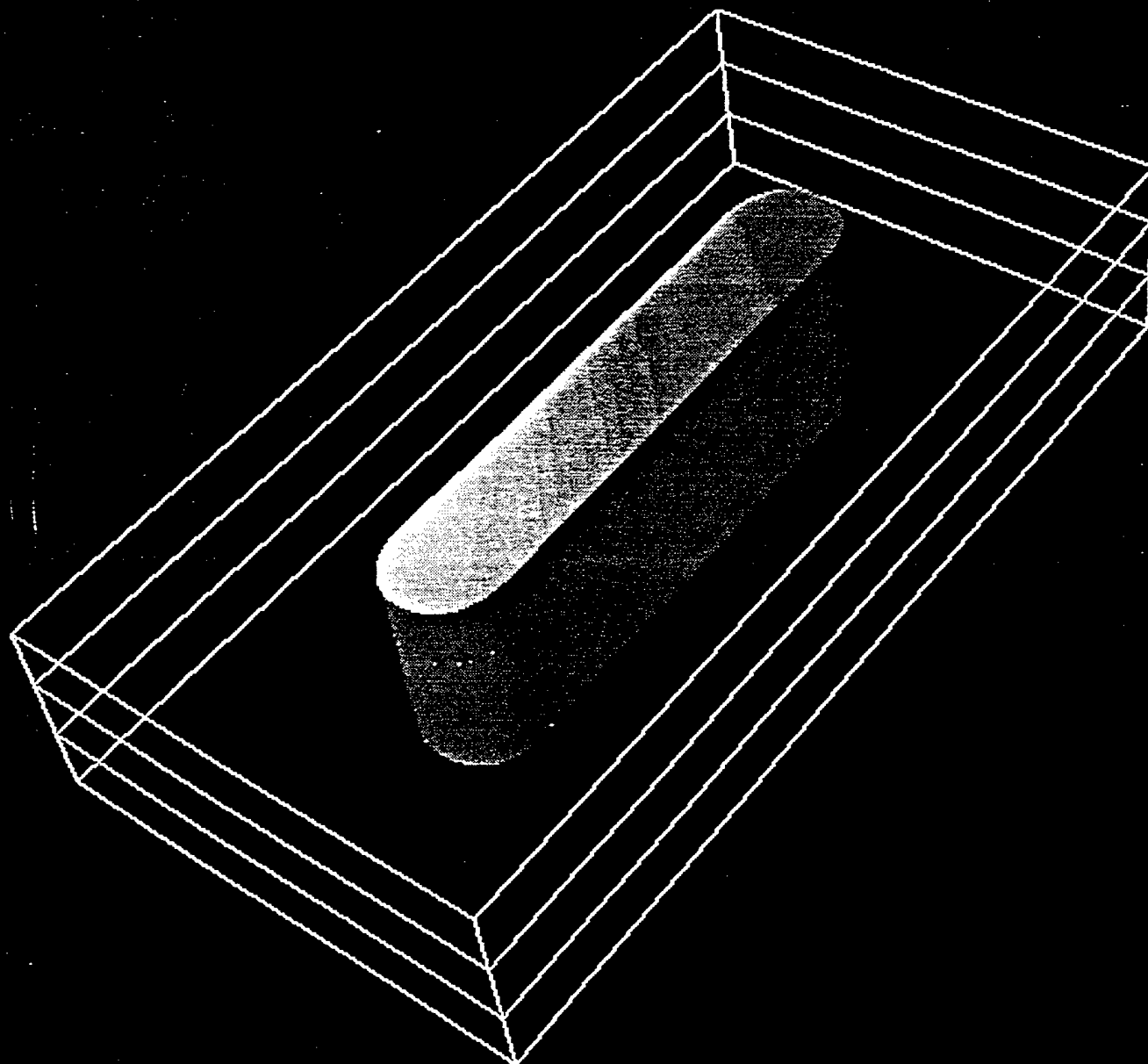
Shortening of L/S of .7u pitch 0def



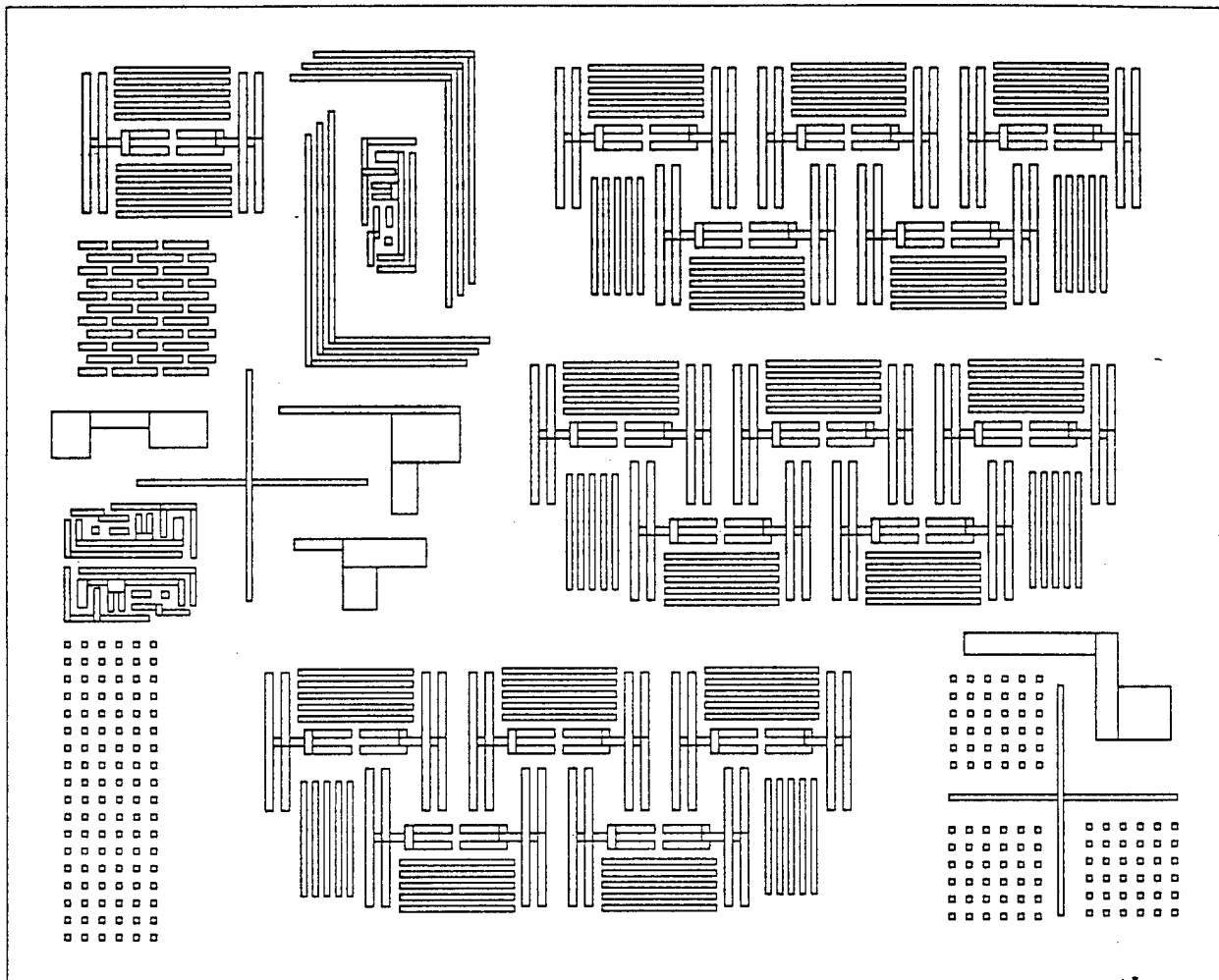
Single Line at .4u 0.6def



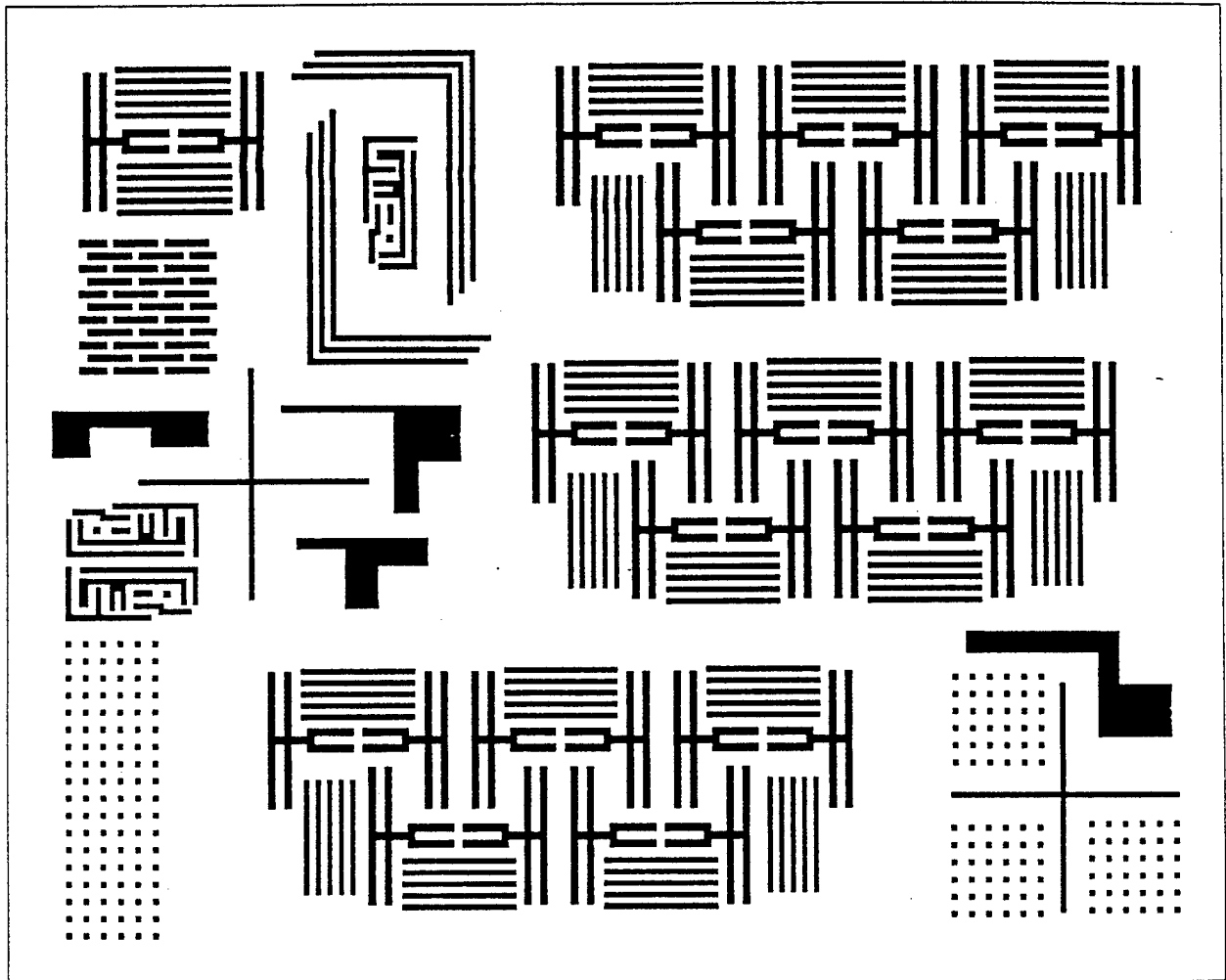
Single Line at .4u 0def



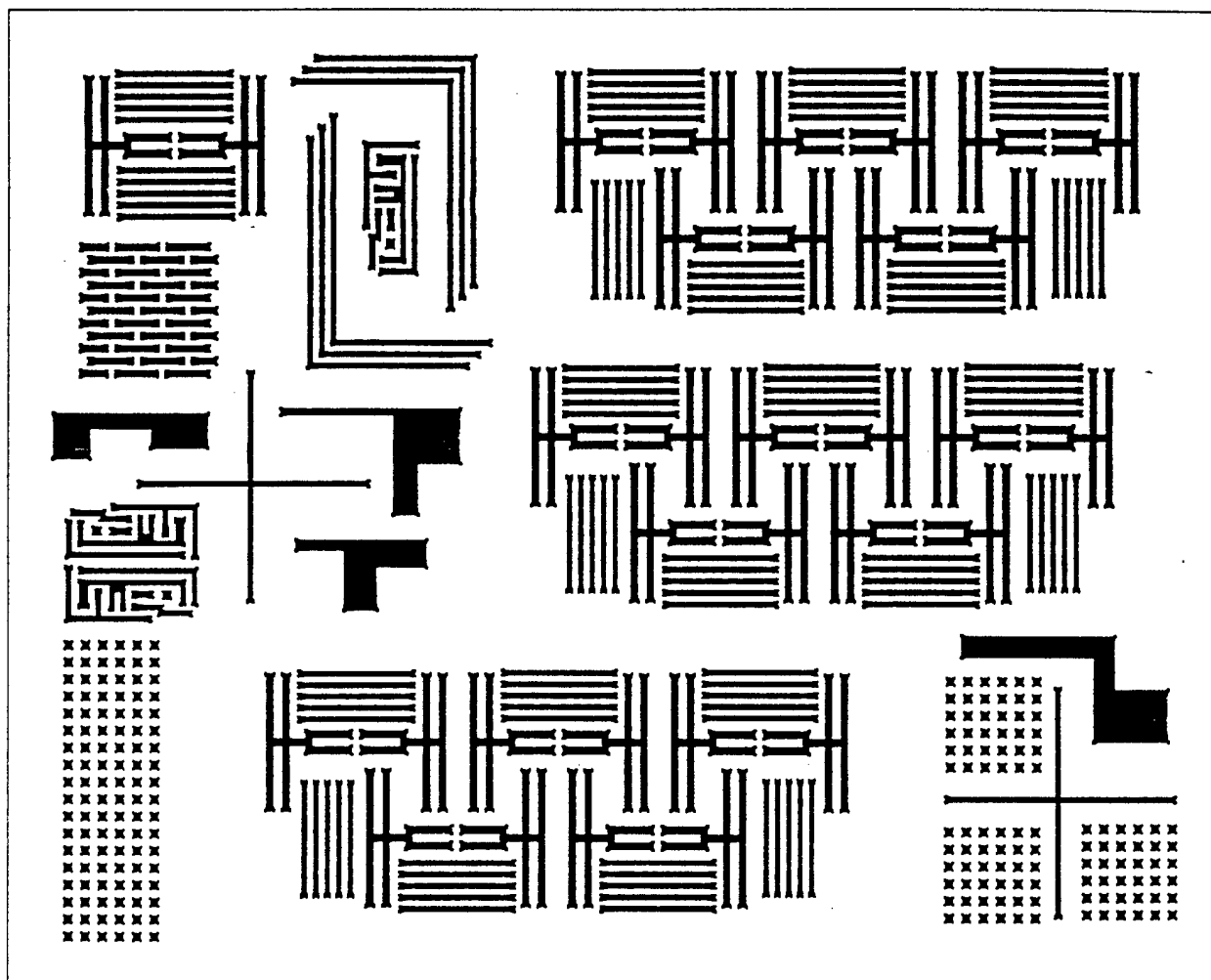
danron5x.input



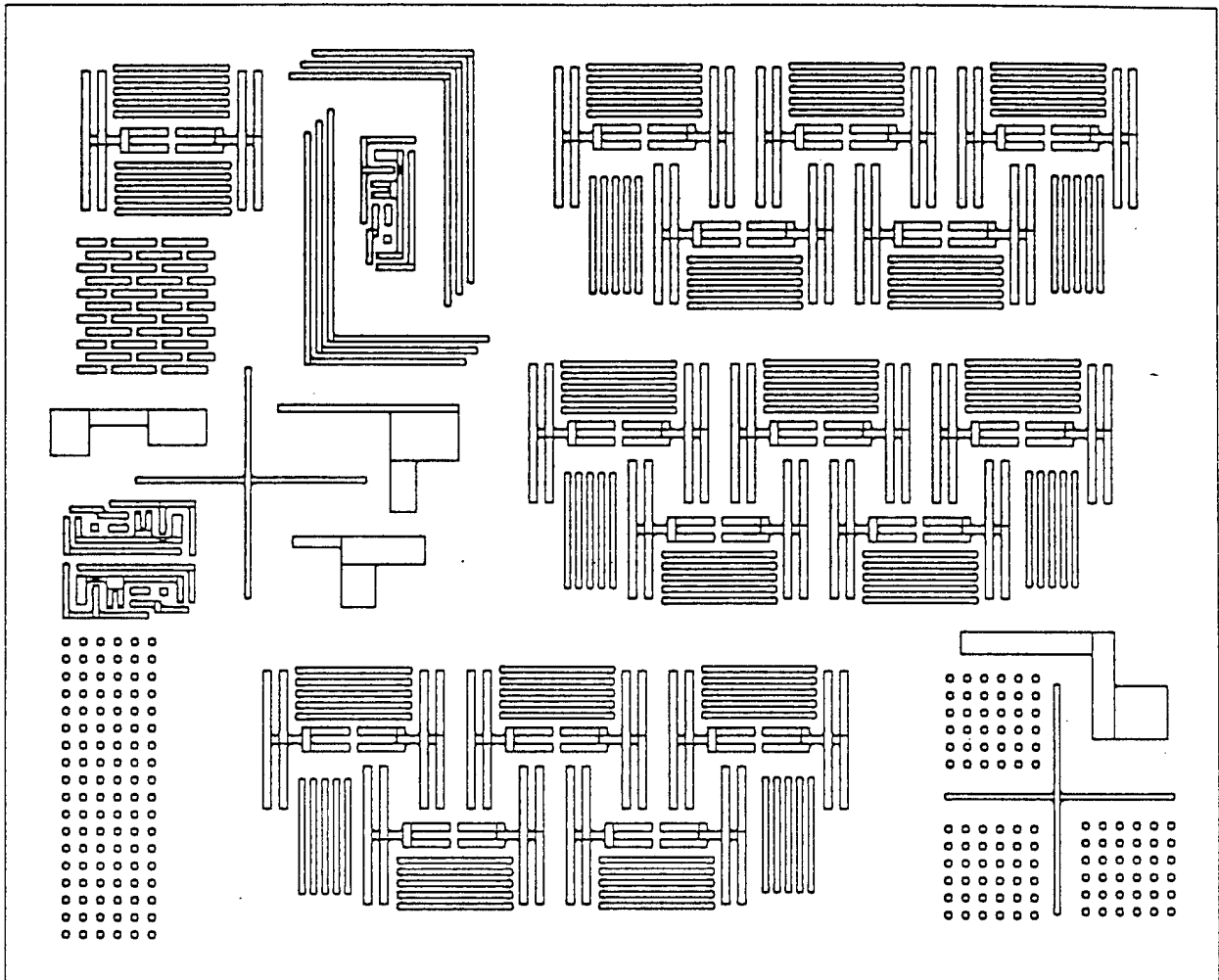
danron5x.iline.faim



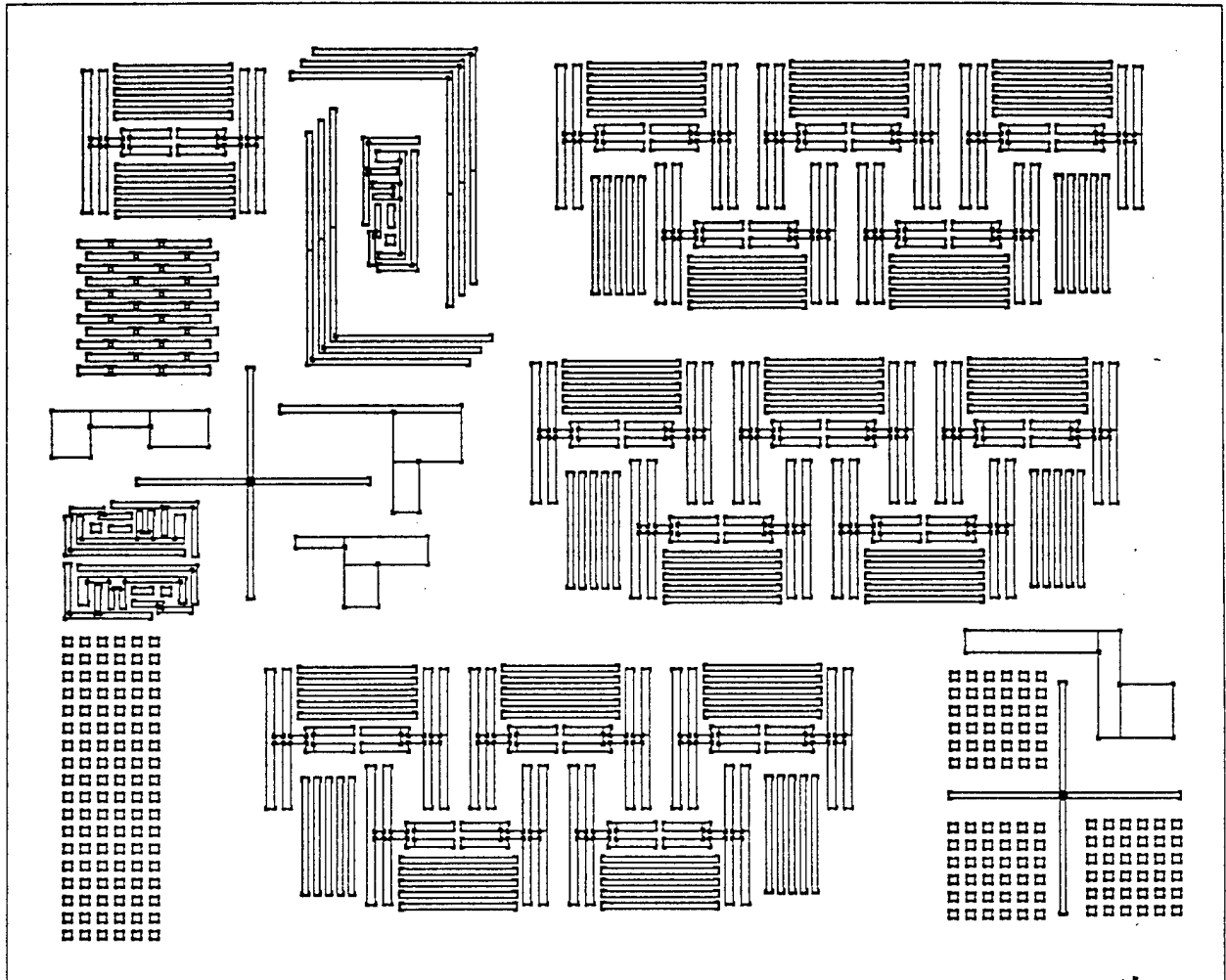
danron5x.iline.optimask



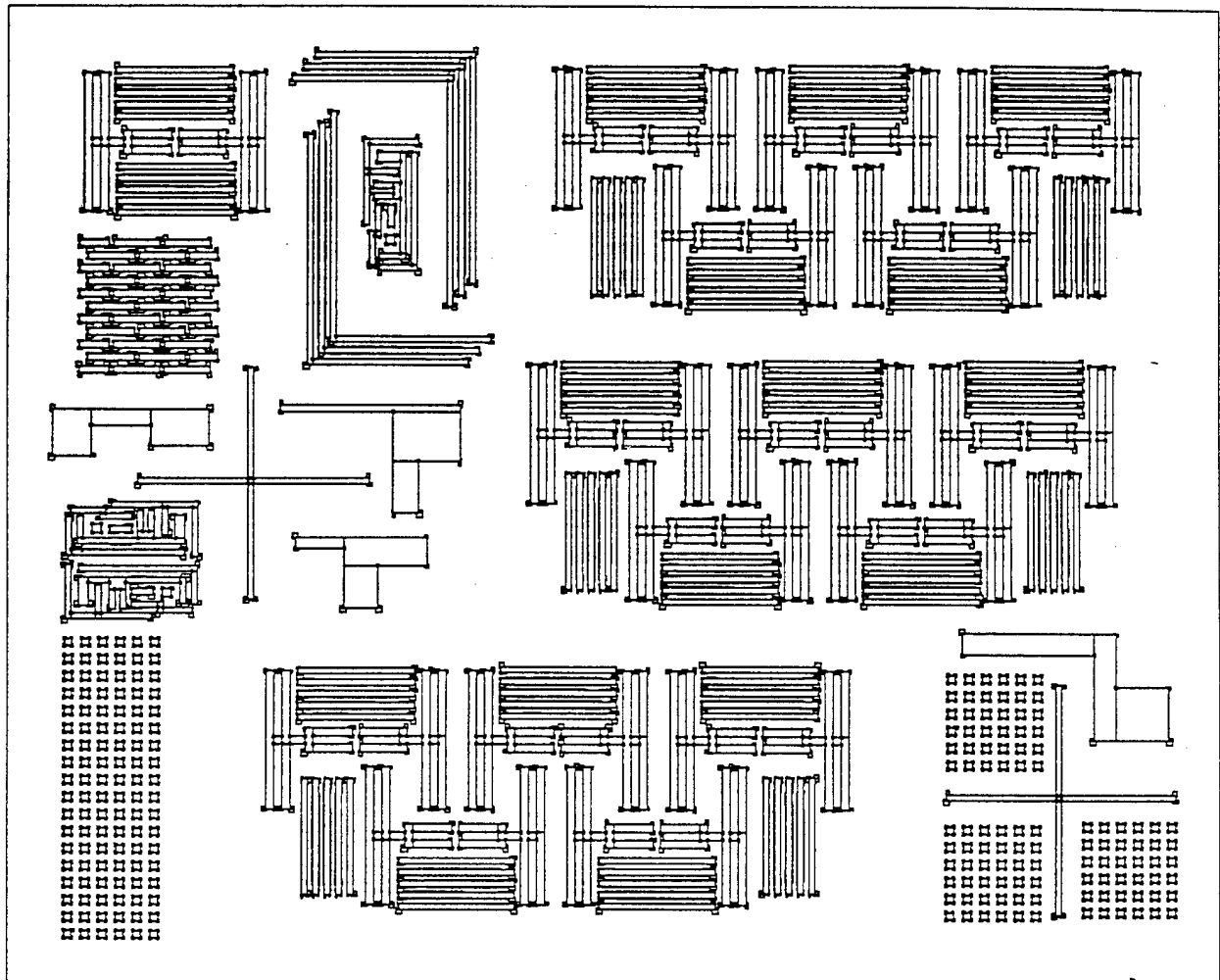
danron5x.iline.faim



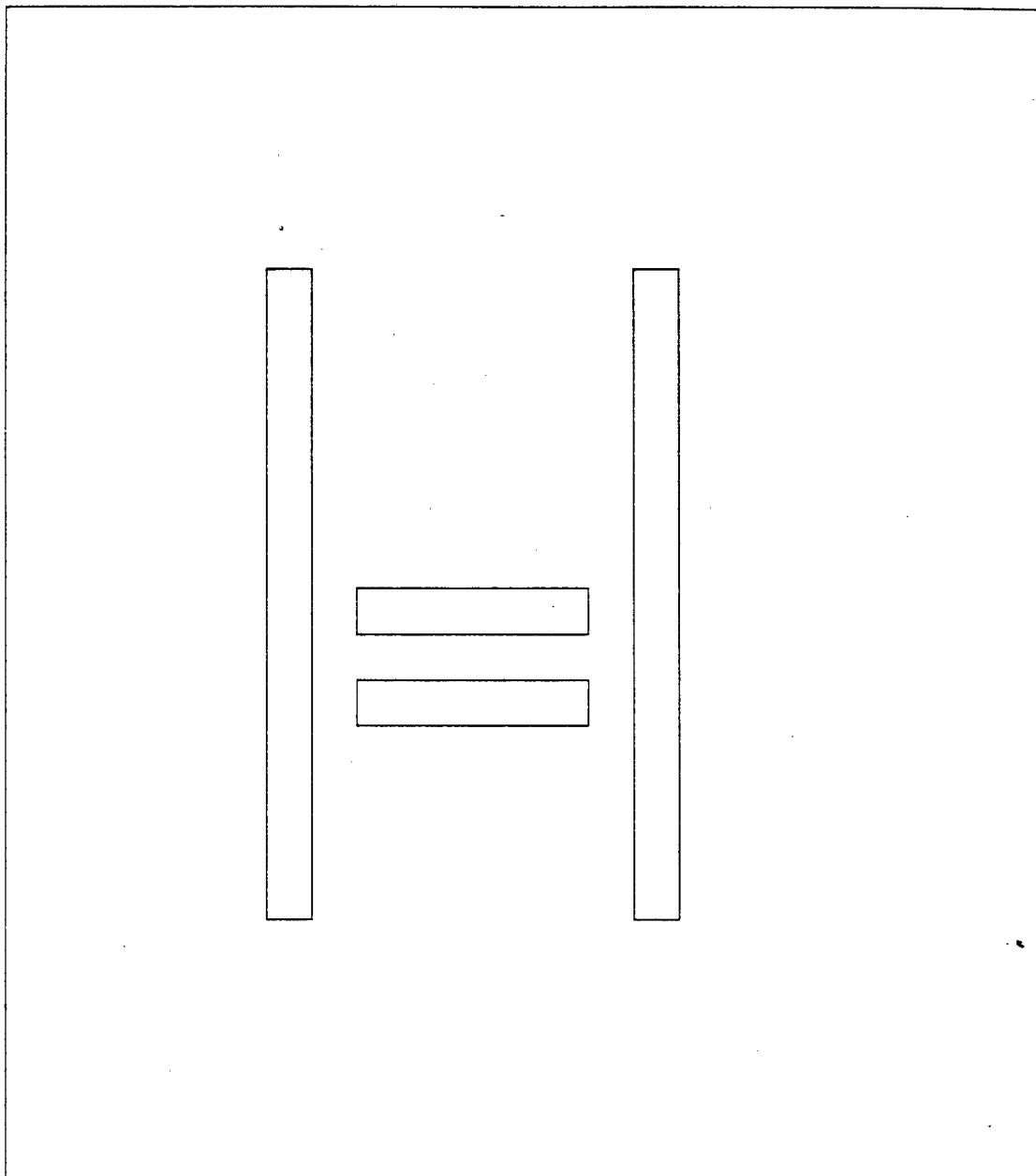
danron5x.optimaskE



danron5x.Eoptimask



test-limop2.faim



Effects of Aliasing Errors on Microlithographic Image Computations

Uwe Hollerbach
Dept. of Manufacturing Engineering
Boston University
Boston, Mass. 02215

Abstract

Recent advances in computational microelectronics have made it possible to compute the images of very large masks. Images of entire masks require many gigabytes of storage. It is therefore desirable to make storage requirements as small as possible. In this paper we investigate the effects of reducing resolution to the point where some aliasing error enters the final answer. Due to the accuracy requirements of the microelectronics industry, we conclude that reducing resolution to improve storage requirements is not feasible.

1. Introduction

Recently, it has become possible to compute quite exactly the images of very large mask patterns used in the microelectronics industry.^{2,3} One effect of this new capability is that the size-limiting factor has become storage rather than computation time. For example, we have computed the image of one layer of an entire SRAM mask. The computation time for this mask was about 24 hours on a 32-node IBM PVS parallel processor, and the storage required for the final answer was approximately 5 Gb. This computation, and others like it, were performed at the minimum resolution required to guarantee that there would be no aliasing in the final answer. To alleviate the storage problem, and also to reduce somewhat further the computation time, it has been suggested that these computations be performed at a lower resolution. Decreasing the resolution by a factor of 2 would save a factor of 4 both in computation time and in storage time, at the cost of accepting some aliasing error in the result. In this paper, we investigate the precise effect of the aliasing error on the answer, with attention given to the particular requirements of the microelectronics industry.

2. Theory

It is well-known^{1,4} that lenses act as spatial filters when creating an image of an object: components of the electromagnetic field with low spatial frequency (slow variation in X

and Y) are passed through, while components with high spatial frequency (fast variation in X and Y) are removed. The coherent transfer function of a simple imaging system (the complex amplitude of a disturbance at the image point (x_i, y_i) due to a unit disturbance at the object point (x_o, y_o) — see figure 1) is given by

$$K(x_o, y_o, x_i, y_i) = \frac{1}{M} \int_0^{\text{NA}} \int_0^{2\pi} e^{ik(\Phi - (x_i + x_o/M)\rho \cos \theta - (y_i + y_o/M)\rho \sin \theta)} \rho d\rho d\theta.$$

Here NA is the so-called numerical aperture. This is a dimensionless measure of the radius of the imaging system, defined as $r_L / \sqrt{r_L^2 + d_i^2}$. M is the magnification factor of the imaging system. ρ and θ are dimensionless polar coordinates in the lens plane. Φ is an aberration function, which contains true lens aberration terms as well as the defocus z_i .

This expression for the coherent transfer function is derived by starting from the Kirchoff theory of diffraction and making the thin-lens approximation and the paraxial approximation. The thin-lens approximation means that the lens system introduces only a phase shift, but no lateral shift, into the electric field as it passes through the lens system. The paraxial approximation means that only rays that are nearly parallel to the optical axis are considered, implying that $\sin \phi \approx \phi$, where ϕ is the angle between the ray and the optical axis. Both of these assumptions are made solely to simplify the presentation; without them, the formulas become more complex, but the conclusions remain equally valid. The computations presented below are in fact made using the more exact formulas as implemented in the FAIM (Fast Aerial Image Model) program.^{2,3}

By scaling the object coordinates as $x'_o = -x_o/M$, $y'_o = -y_o/M$, the coherent transfer function may be written as

$$K(x_i - x'_o, y_i - y'_o) = \frac{1}{M} \int_0^{\text{NA}} \int_0^{2\pi} e^{ik(\Phi - (x_i - x'_o)\rho \cos \theta - (y_i - y'_o)\rho \sin \theta)} \rho d\rho d\theta$$

The electric field in the image space is then given by

$$\begin{aligned} E(x_i, y_i) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(x'_o, y'_o) K(x'_o, y'_o, x_i, y_i) dx'_o dy'_o \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(x'_o, y'_o) K(x_i - x'_o, y_i - y'_o) dx'_o dy'_o, \end{aligned}$$

where F is the function which defines the electric field as a function of position in the object plane. (In general this would be a function of position in the three-dimensional object space, but for microlithographic applications it is sufficient to consider planar objects only.) The expression for the electric field is a convolution, which may be evaluated very simply by Fourier-transforming:

$$\hat{E}(f_x, f_y) = \hat{F}(f_x, f_y) \hat{K}(f_x, f_y)$$

Because the lens system has a finite radius, the Fourier transform \hat{K} of the coherent transfer function is zero beyond a certain radius in spatial frequency space (f_x, f_y): specifically, the cutoff is at NA/λ . According to the Nyquist theorem, it is therefore necessary to use a spatial grid with a spacing of $\lambda/2NA$ or finer in order to properly resolve all frequency components that are present in the electric field in the image space. In order to compute the intensity in the image space, the electric field is squared. This squaring doubles the highest spatial frequency that is present, so that in order to adequately resolve all components of the image intensity, it is necessary to use a spatial grid with a spacing of $\lambda/4NA$ or smaller. If a coarser grid is used, there will be aliasing effects.

3. Practice

In microlithographic applications, the object being imaged is the mask, which is described by a piecewise-constant discontinuous function. Errors due to aliasing may be regarded as either locally changing the value of the mask function, or as changing the positions of the discontinuities. Of these two types of error, the second is much more important for microlithographic applications. The discontinuities mark the pattern which is to be printed, and changes in position of even a few tens of nanometers (a few percent of the nominal linewidth) can mean the difference between a working and a non-working integrated circuit.

Figure 2 shows a small sample pattern, about $14\mu m \times 10\mu m$, that is representative of mask patterns actually used in the microelectronics industry. The lines and the spaces between them are $0.35\mu m$ in size. The wavelength used to image this sample pattern is $0.365\mu m$, and the NA is 0.55. These values are typical for current (1996) computer chip production. Figure 3 shows the image computed at a resolution of $\lambda/16NA$. This image is essentially exact; there are no aliasing errors present. The intensity contour levels shown are 0.1, 0.3, and 0.5. Note that the curves are quite smooth, except where features interact. This smoothness is expected: the filtering action of the lens system tends to smooth out the pattern being imaged. In marked contrast, the image shown in figure 4 shows significant ripples in the contour levels. The only difference between the two images in figures 3 and 4 is that the latter was computed at a resolution of $\lambda/2NA$. The ripples are unphysical and are solely due to the aliasing caused by the inadequate resolution. Figure 5 shows the difference between the two images, overlaid with the actual mask pattern. The shaded regions indicate where the magnitude of the difference exceeds 0.2. From this figure, it is quite clear that the aliasing errors which have been introduced into the low-resolution image significantly change the effective position of the pattern edges. Finally, in figure 6 the maximum error due to aliasing is plotted as a function of the relative resolution for several different images. A relative resolution of 1 means that the grid size is $\lambda/4NA$. In all cases, the maximum error at a resolution of $\lambda/2NA$ (a relative resolution of 1/2) is quite large, of the order of 0.3 or more. Since the maximum intensity is between 1 and 1.2 in all cases, an error of this magnitude is very significant. In order to achieve a maximum error that is smaller than 0.01, it is necessary to have a relative resolution of 0.9 or higher. For relative resolutions larger than or equal to 1, the behavior of the error depends on whether

the image was computed assuming a periodic mask pattern or an aperiodic pattern. The error of $\sim 3 \times 10^{-5}$ in the aperiodic case is due to the way in which the edges of the overall domain are treated in the computation. Because this edge effect is quite small compared with the aliasing errors which are present at relative resolutions less than 1, and because it occurs well away from the mask patterns which are of physical interest, it may safely be ignored. If a periodic pattern is assumed, there is no aliasing error. There is a very small difference of about 10^{-9} , not shown in the figure, but this is caused by roundoff errors in writing the computed intensity field to a file, not by any aliasing error during the computations.

4. Conclusion

To summarize, it is evident that aliasing errors are present in image intensity computations made at resolutions of less than $\lambda/4NA$. Due to the particular nature of the object being imaged, and due to the specific needs of the microelectronics industry, the magnitude of these aliasing errors is unacceptably high at resolutions near $\lambda/2NA$. Since it is necessary to go almost to a resolution of $\lambda/4NA$ in order to achieve an error below 0.01, it must be concluded that it is not possible to achieve significant time or storage savings by sacrificing accuracy in resolution.

5. Acknowledgements

This work was supported in part by the Advanced Research Projects Agency, by the Air Force Office of Scientific Research, and by the National Institute of Standards and Technologies.

6. References

- 1) M. Born and E. Wolf, *Principles of Optics*, 6th ed. (Pergamon, New York, 1980).
- 2) D. Cole, E. Barouch, U. Hollerbach, and S. A. Orszag, "Calculation of Aerial Images for High Numerical Aperture Projection Systems", *J. Vac. Sci. Technol.* bf B10, 3037 (1992).
- 3) D. Cole, E. Barouch, U. Hollerbach, and S. A. Orszag, "Derivation and Simulation of Higher Numerical Aperture Scalar Aerial Images", *Jap. J. Appl. Phys.* bf 31 (48) 4110 (1992).
- 4) J. W. Goodman, *Introduction to Fourier Optics* (McGraw-Hill, 1968).

7. Figure Captions

Fig. 1. Diagram of a simple thin-lens optical system. The "thin-lens" approximation means that refraction effects at the front and back surfaces of the lens are not considered separately, but are lumped together into a single change of direction at the central plane of the lens.

Fig. 2. A small sample mask pattern. The lines to be printed are $0.35\mu m$ thick.

Fig. 3. Contour lines of the correctly computed image of the mask in figure 2. The intensity levels shown are 0.1, 0.3, and 0.5. Notice that the contour lines are quite smooth, except where features interact.

Fig. 4. Contour lines of the under-resolved, incorrectly computed image of the mask in figure 2. The intensity levels shown are again 0.1, 0.3, and 0.5. In contrast to figure 3, the contour lines are wavy even where no waviness is expected.

Fig. 5. The difference between the correct and incorrect intensity fields. The gray blotches indicate those regions where the difference exceeds 0.2. Notice that the blotches lie primarily near or on the nominal mask edges.

Fig. 6. The maximum aliasing error as a function of the relative resolution for several different images. The first two sets of points are for the periodic and non-periodic versions of the mask shown in figure 2, while the last two are for different mask patterns which are not shown here.

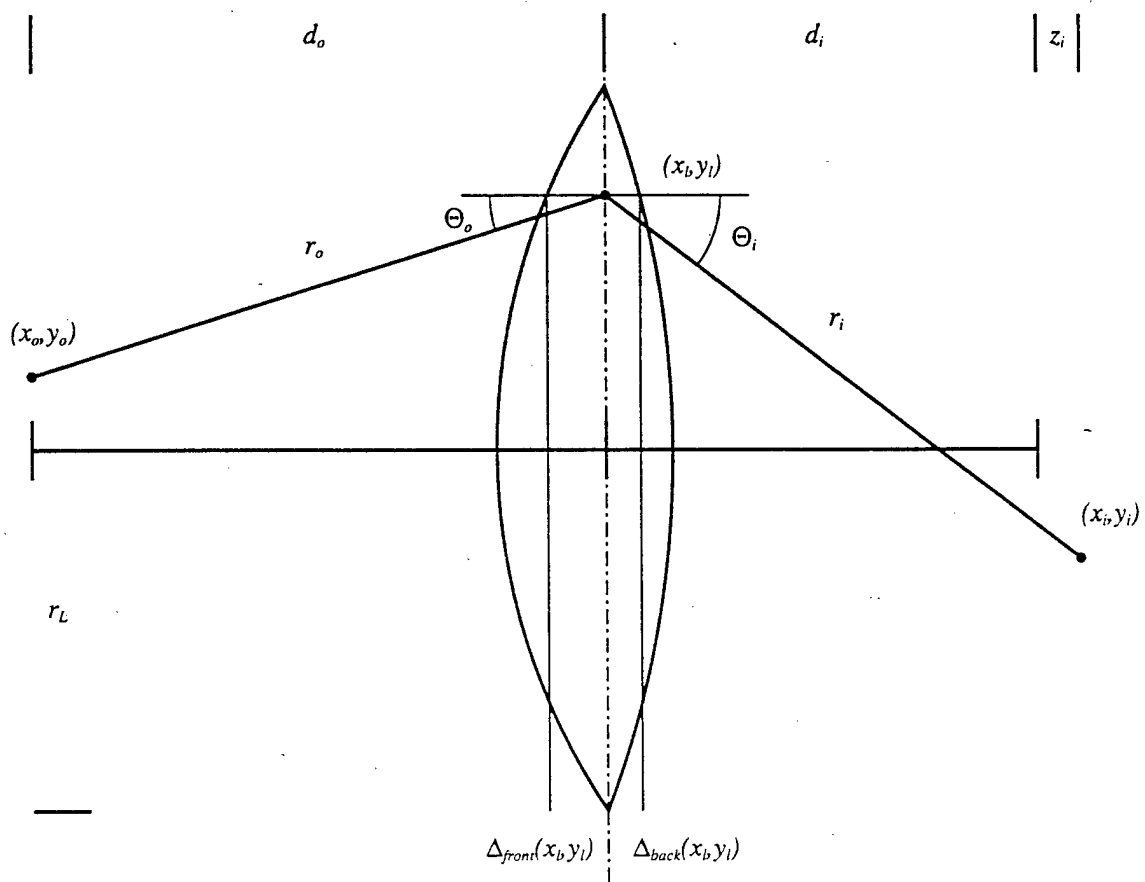


Figure 1

Sample Mask Pattern

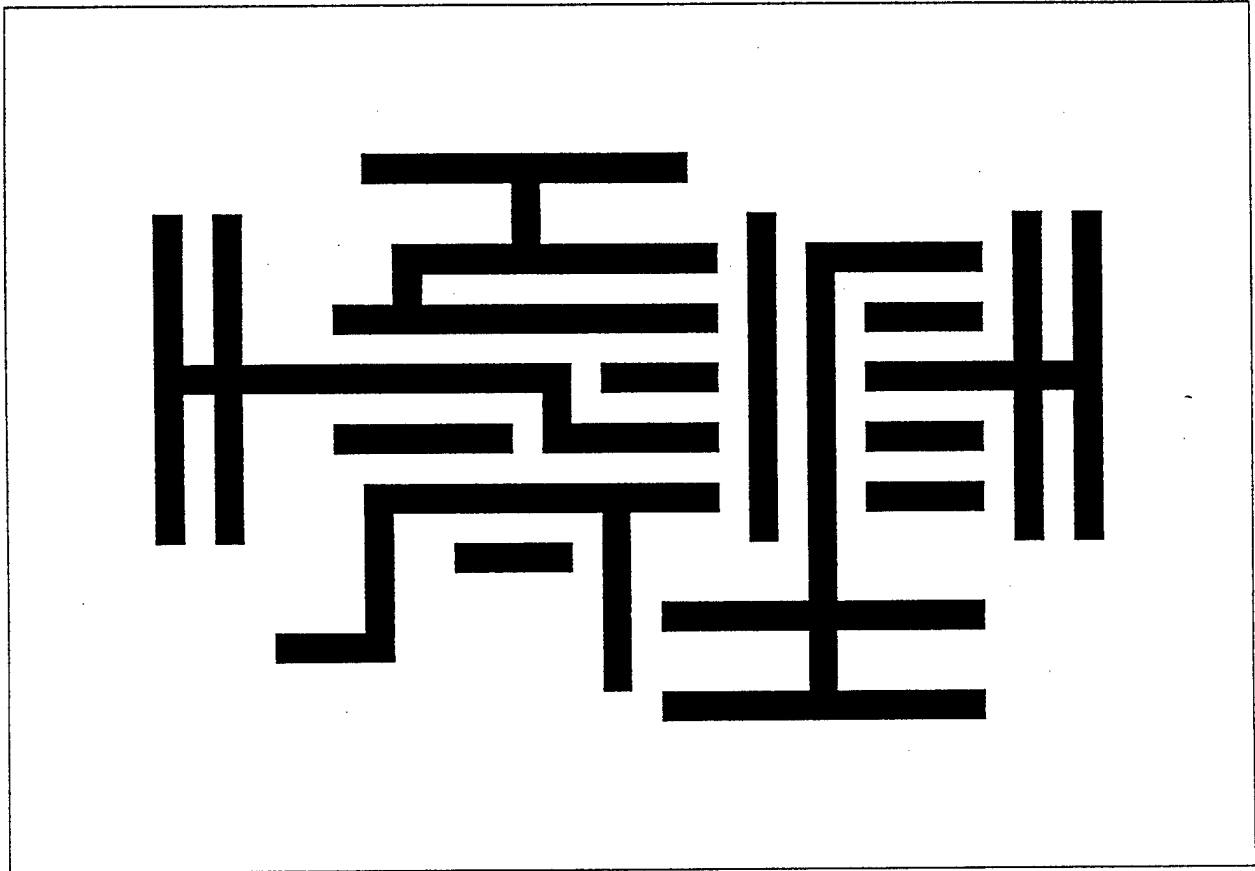


Figure 2

Sample Mask Imaged at High Resolution

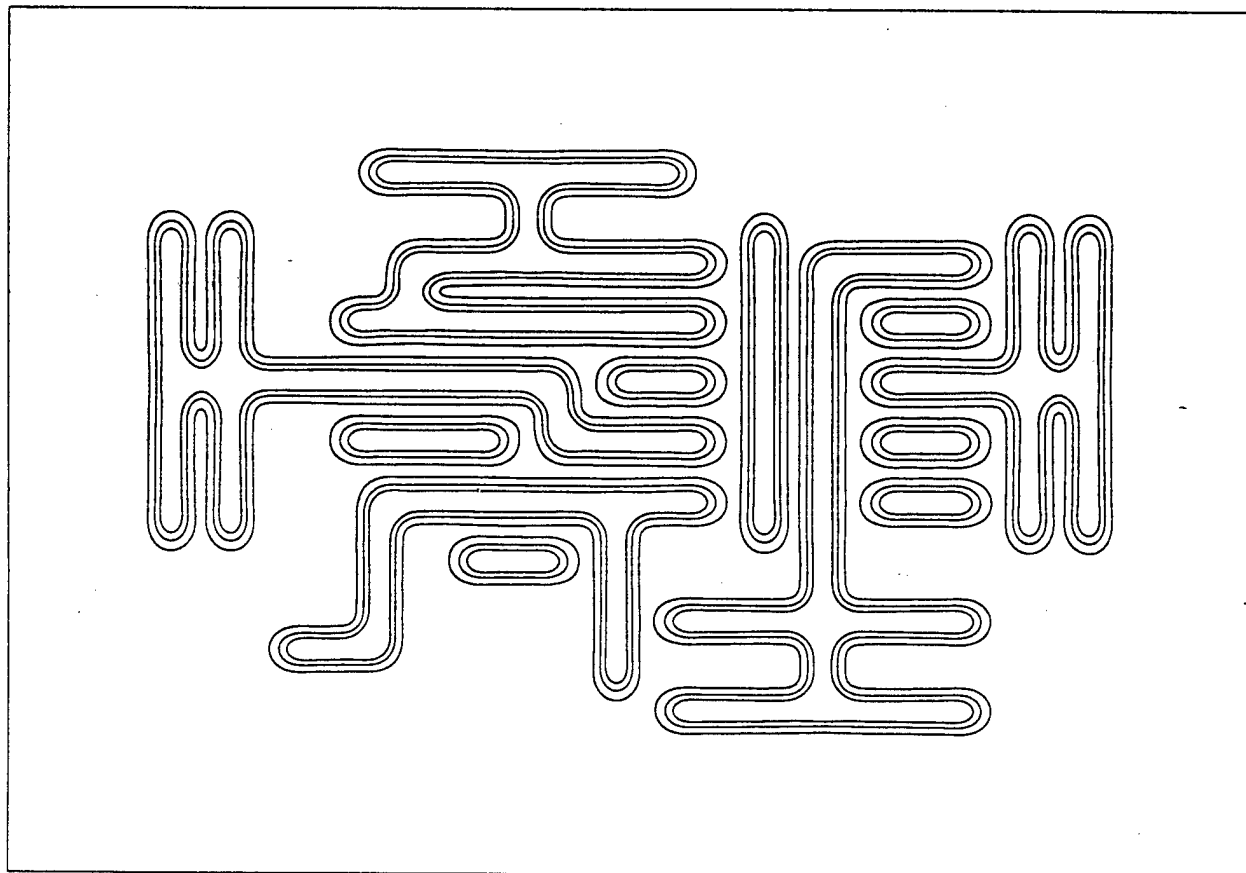


Figure 3

Sample Mask Imaged at Low Resolution

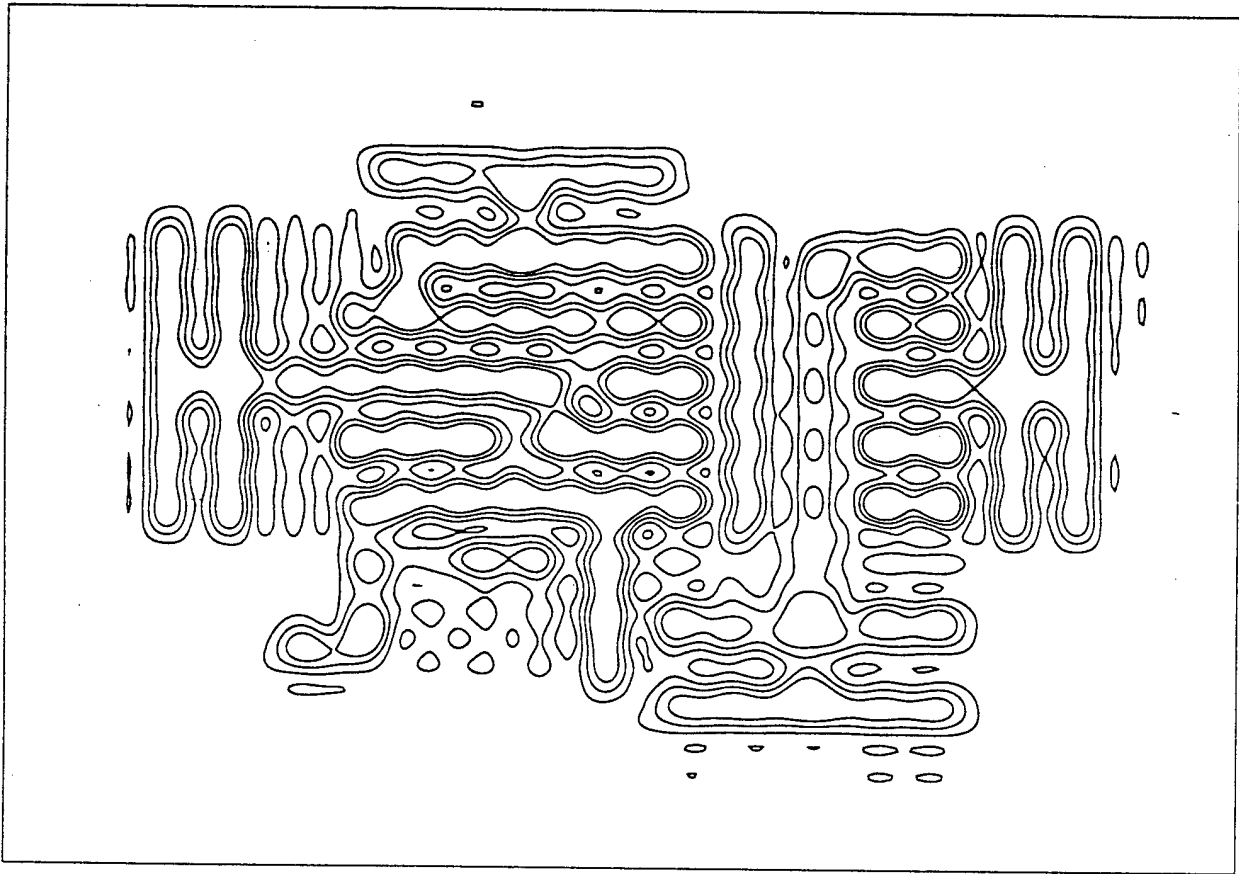


Figure 4

Difference Between High- and Low-Resolution Images

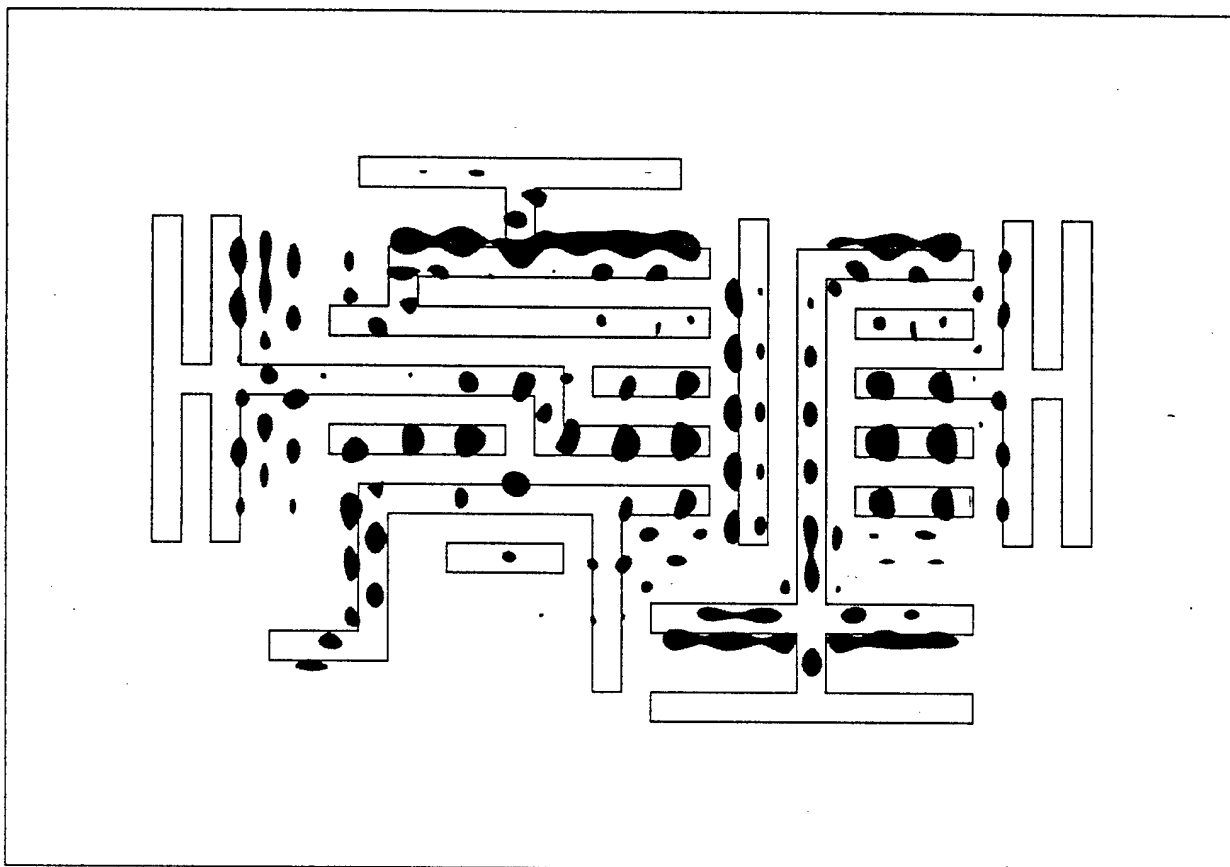


Figure 5

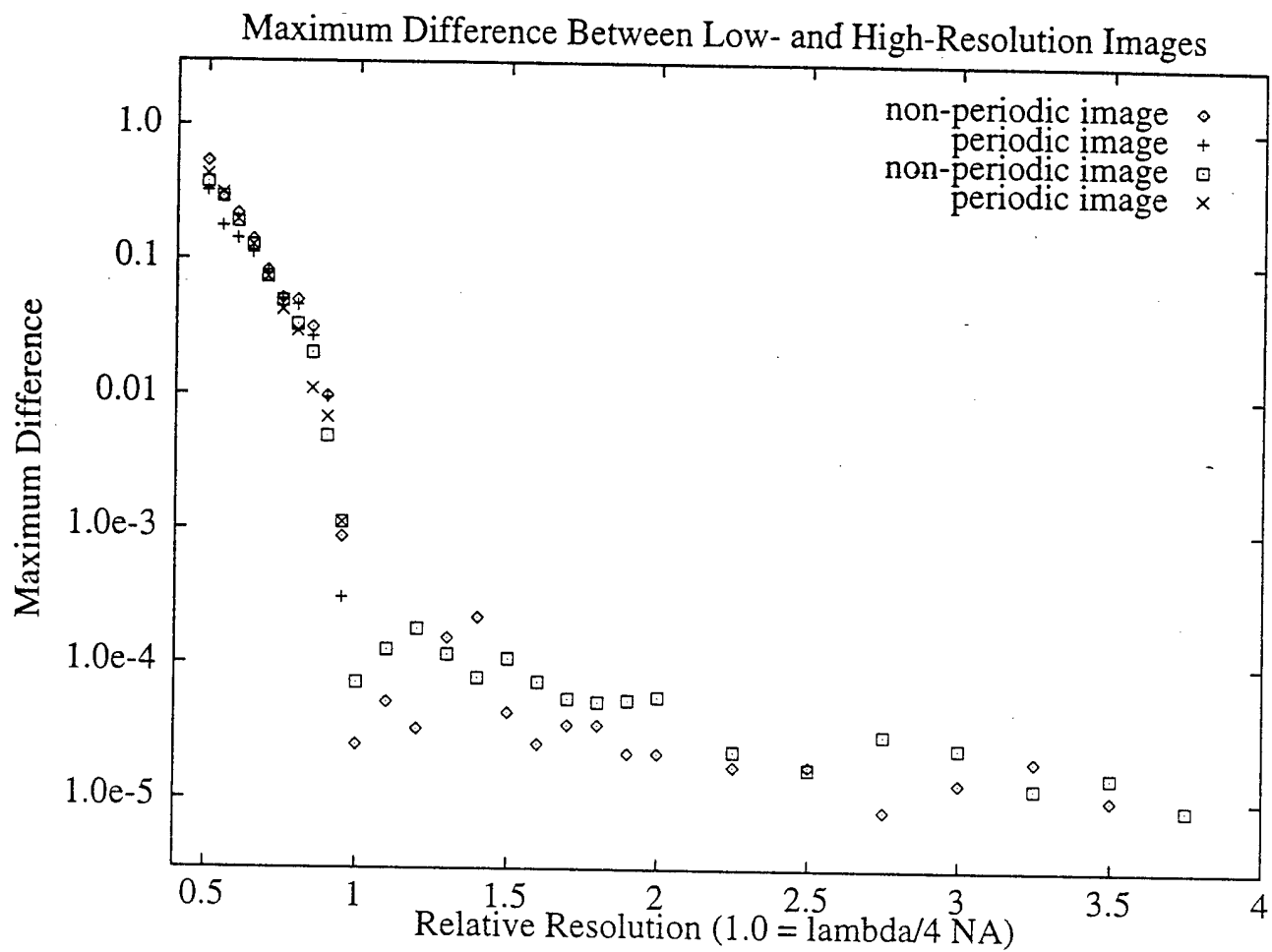


Figure 6

Influence of mask topography on mask image

Vadim Borue

May 15, 1996

Phase-shifting masks have noticeable width. They are fabricated with the layer of quartz on the top (or bottom) of a thin chrome layer. To obtain phase shift φ quartz is etched back on the depth

$$d = \frac{\lambda\varphi}{2\pi(n-1)} \quad (1)$$

where λ is wave-length and n is quartz refractive index. It is natural to expect that finite mask thickness not only leads to extra phase but also will lead to a certain corrections in mask transfer functions.

1 Basic formulas

1.1 Plane waves representation

We use only scalar diffraction approximation in Kirchoff approximation. We define the basic wave-number $k_0 = 2\pi/\lambda$. The wave-number is represented as (\mathbf{k}, k_z) and $\mathbf{k} = (p, q)$ is a two-dimensional wave-number ($k_z^2 + \mathbf{k}^2 = k_0^2$). If the Fourier transform of the field $u(x, y, z)$ (satisfying Helmholtz equation) is known at plane z' then the Fourier transform of the same field in the half plane $z > z'$ without extra sources is

$$\hat{u}(\mathbf{k}, z) = \exp[-ik_z(z - z')] \hat{u}(\mathbf{k}, z') \quad (2)$$

It is also assumed that $k_z = (k_0^2 - \mathbf{k}^2)^{1/2}$ the sign corresponds to correct direction of propagation. Complex k_z corresponds to evanescent waves.

If numerical aperture of the optical or illuminating system is small then we may assume that $|\mathbf{k}| \ll k_z$ and $k_z \approx k_0 - |\mathbf{k}|^2/(2k_0)$. This approximation is known as Fresnel approximation or parabolic approximation. Indeed, returning to physical space and assuming that $|z - z'| = d$

$$u(\mathbf{x}, z) = i \frac{\exp(-ik_0 d)}{\lambda d} \int u(\mathbf{x}', z') \exp \left[-i \frac{k_0}{2d} (\mathbf{x} - \mathbf{x}')^2 \right] d^2 x' \quad (3)$$

and this is obviously Fresnel representation.

In the Fraunhofer limit

$$u(\mathbf{x}, z) = i \frac{\exp(-ik_0 R_0)}{\lambda d} \int u(\mathbf{x}', z') \exp \left[ik_0 \frac{\mathbf{x} \cdot \mathbf{x}'}{d} \right] d^2 x' \quad (4)$$

where $R_0 = (d^2 + \mathbf{x}^2)^{1/2}$. That is equivalent to

$$u(\mathbf{x}, z) = i \frac{\exp(-ik_0 R_0)}{\lambda d} \hat{u}(k_0 \frac{\mathbf{x}}{d}, z') \quad (5)$$

That is an image is a Fourier transform of initial distribution taken at wave-numbers corresponding to angles of diffraction. On the other hand if $R_0 \approx d$ this formula may be inverted

$$\hat{u}(\mathbf{k}, z) = i \frac{\exp(-ik_0 d)}{2\pi} u(d \frac{\mathbf{k}}{k_0}, z') \quad (6)$$

1.2 Object illumination

We assume that Kohler or critical illumination is used. In this case it follows that mutual coherence may be represented as

$$J(x - y) = \int d^2 k F(k) \exp^{ik(x-y)} \quad (7)$$

where $F(k)$ is an amplitude of a coherent plane wave. Moreover $F(k)$ equals to the form of the pupil function of an apperture (with the numerical aperture N_c) used in the condenser. For example, if a condenser aperture is circular.

$$F(k) = \theta(N_c^2 - |k|^2/k_0^2) \quad (8)$$

Thus to obtain an image from this illumination we need to find an image from each coherent wave and then find the incoherent sum of the waves with corresponding amplitudes.

1.3 Optical system

The lens represents a low-pass filter with lens transfer function $K(x_o, x_i)$, where x_o and x_i coordinates in object and image spaces. If u_o and u_i fields in object and image spaces, respectively. Then

$$u_i(x_i) = \int K(x_o, x_i) u_o(x_o) d^2 x \quad (9)$$

The general form of transfer function is

$$K(k_o, k_i) \approx K_n e^{i \frac{2\pi}{\lambda} T(k_o, k_i)} \delta(k_o + M k_i) \theta(N_a^2 - k_i^2) \quad (10)$$

where k_o and k_i are wave-numbers in object and image spaces, M is angular magnification (or image space reduction). $T(k_o, k_i)$ is the angle characteristic of the lens that should include both Gaussian approximation and aberrations. N_a is numerical apperture of the lens. Final image field is

$$F_i(k_i) = \int F_o(k_o) K(k_o, k_i) d^2 k_o \quad (11)$$

All the imaging properties of the lens depends in this representation on the angle characteristic of the lens.

It is possible to use specific $T(k_o, k_i)$ directly, but usually z_i and z_o are chosen the way that these planes are imaged in each other within Gaussian optics. In this case

$$K(k_o, k_i) \approx K_n e^{i \frac{2\pi}{\lambda} m_i z_i} \delta(k_o + M k_i) \theta(N_a^2 - k_i^2) \quad (12)$$

and z_i is defocus.

1.4 Object transfer function

The monochromatic wave with wave-number k passes trough the object and scater. The amplitude after the object is

$$F_o(k_o) = \int F(k) O(k, k_o) d^2 k \quad (13)$$

where $O(k, k_o)$ object transfer function. In case when $O(x, y) = \delta(x - y) O(x)$ (infinitely thin object) $O(k, k_o) = O(k - k_o)$ and $O(k)$ is Fourier transform

of the object. Therefore object effectively performs convolution in Fourier space. Only small k_0 are imaged in the lens.

Rigorously to solve the problem of thick mask one should solve Maxwell equations at the mask for different incident waves and then decompose the solution in Fourier series which are imaged by lens. Only slowly varying functions on the mask are imaged. This is because of low numerical aperture N_a/M at object side. The most rough estimate of the influence of the thickness of object can be obtained by noticing that by passing through object wave get the phase

$$\exp(i \frac{2\pi}{\lambda} (1 - k^2)^{1/2} d(n - 1)) = \exp(i\varphi(1 - k^2)^{1/2}) \quad (14)$$

Using the fact that $k_i = k_o/M$ we see that this equivalent to introducing defocus $z_o = d(n - 1)/M^2$.

Simplest method of implementation is to replace Fourier transform of object by

$$\exp(i\varphi(1 - k^2)^{1/2}) O(k) \quad (15)$$

where φ is a phase shift of a given object.

Therefore, we may see that first very approximate correction to the mask thickness is an introduction of defocus level

$$z_o = \frac{d(n - 1)}{M^2} \quad (16)$$

This can be interpreted as if imaging of the mask occurs the way that holes without phase-shift are imaged as they are. The hole with phase-shift are imaged the way as if they are defocused by the amount z_o . The reason why these holes are not equivalent is that side-walls of the hole have nearly zero amplitude of the field on them [3]. These follows mostly because the angle of incidence is very small and the hole plays the role of perfectly reflecting waveguide. Then in this case we can use Kirchhoff representation using the interface between air and quartz instead of aperture.

Obviously that is very crude estimate. The only way I see to improve it is to solve the whole problem. This way is obviously impractical. May be an alternative way is to tabulate various defocuses for various form and materials and use these as the correction.

First it should be verified that mask thickness may be taken care of by introducing defocus. One more reason why this can be so is that only very

smoothly varying transfer functions on the hole can be imaged by small numerical aperture (on the object side) optical system. In this case we may expand corrections in Taylor series. There should not be linear terms because of symmetry. Quadratic terms of the expansion are essentially equivalent to introducing small defocus. Indeed quadratic variations in phase in final integral are equivalent to quadratic variations in space for aberration function and that means the presence of defocus.

From the experimental point of view the presence of defocus was underlined in [4], [5]. Although it was not checked there that the hole change of the image can be described by introducing defocus this seems plausible judging from the data of [4]. Also it was noticed in [3] that the influence of thickness of mask decrease dramatically with increase of magnification M . Indeed it should decrease quadratically.

The defocus was observed in [5] to depend slightly on the width of the hole and on whether the mask is binary or attenuated phase-shifting. So far the only conclusion that can be drawn is that accounting for the width of the mask by introducing additional defocus may be convenient parametrization of the effect of the mask. This parametrization is purely empirical and requires thorough testing with rigorous codes such as EMFLEX or TEMPEST. I think that we need either write this type of codes ourselves or buy them.

References

- [1] M. Born, E. Wolf, Principles of optics, Pergamon, New York, 1985.
- [2] D.C. Cole, E. Barouch, U. Hollerbach, S.A. Orszag, "Derivation and simulation of higher numerical resolution aperture scalar aerial images", Jpn. J. Appl. Phys. **31**, 4110-4119 (1992).
- [3] G. Wojcik, J. Mould, R. Ferguson, R. Martino, K.K. Low, "Some image modelling issues for i-line, 5x phase shifting masks." SPIE, **2197**, 455-465, (1994).
- [4] A.K. Wong, R. Ferguson, R.M. Martino, A.R. Neureuther, "Focus shift and process latitude contact holes on attenuated phase-shift masks", SPIE, **2440**, 472-479 (1995).

- [5] A.K. Wong, R. Ferguson, A.R. Neureuther, "Phase-shifter edge effects on attenuated phase-shifting mask image quality", SPIE 2197, 122-129 (1994).

Optimization of Stepper Parameters and their influence on OPC

Rakesh R. Vallishayee and Steven A. Orszag

Princeton University, Fluid Dynamics Research Center,
Princeton, NJ 08544.

Eytan Barouch

Boston University, Department of Manufacturing Engineering, Boston, MA 02144.

ABSTRACT

An algorithm for the optimisation of stepper parameters has been designed and implemented. The cost function used in this optimisation is the contrast. The aerial image is computed using the computer code FAIM.^{1,2} First, the contrast of the image is calculated and the derivatives of the contrast with respect to the stepper parameters are evaluated. The computational cost of these calculations is only slightly more than that of one aerial image simulation. A conjugate gradient type algorithm is then used to obtain the minimum of the contrast.

1. INTRODUCTION

Stepper conditions have a significant effect on mask printability. It is therefore very important to optimize the stepper conditions to yield the best possible stepper performance. Traditionally this is done through trial and error. Here we report the development and implementation of an algorithm that shortens this procedure by optimising the illuminator design and numerical aperture averaging over several defocus values.

We first compute the aerial image of the given mask using our Fast Aerial Image Module code, FAIM. Then we calculate the contrast associated with the aerial image and the derivative of the contrast with respect to the stepper parameters to be optimised. Finally the contrast is minimised by taking an iterative step in the appropriate direction determined using a conjugate gradient type algorithm.

In this paper we will first describe the procedure for the optimization of the stepper parameters (cost, derivatives of cost with respect to the optimisation parameters, constraints and convergence criteria) at fixed defocus. Then we will discuss the optimisation of stepper parameters over a specified defocus range for the purpose of doing optical proximity corrections (OPC). We will show examples of how the optimisation procedure works on some current mask designs.

2. OPTIMIZATION PROCEDURE

2.1 Cost and it's derivatives

We define I_0 to be the ideal intensity pattern that we wish to obtain on the wafer or image plane. For example, with a dark field mask the value of I_0 is 1 on portions of the wafer plane that we wish to expose and 0 on portions that we do not wish to expose. We define I to be the intensity pattern obtained on the wafer plane using the FAIM simulation. Then we can write the contrast, which is the cost to be minimised, as,

$$C = \int_{\text{wafer}} (I - I_0)^2 dA \quad (1)$$

where C is the contrast and dA refers to an area element of the wafer. The objective of the optimization is to minimize C so that I is as close to I_0 as possible. I_0 is a constant during the optimization and I is a variable which varies when the stepper parameters are changed. Let us refer to the stepper parameters in general as p_1, p_2, \dots, p_n . Then the variation of the cost with respect to the stepper parameters can be written as

$$\begin{aligned} \delta C &= 2 \left[\int_{\text{wafer}} (I - I_0) \frac{\partial I}{\partial p_1} dA \right] \delta p_1 + \dots + 2 \left[\int_{\text{wafer}} (I - I_0) \frac{\partial I}{\partial p_n} dA \right] \delta p_n \\ &= \frac{\partial C}{\partial p_1} \delta p_1 + \dots + \frac{\partial C}{\partial p_n} \delta p_n \end{aligned} \quad (2)$$

where $\partial C / \partial p_r$ is the partial derivative of C with respect to p_r . We see from (2) that if the signs of δp_r 's are opposite to that of $\partial C / \partial p_r$'s, then δC will always be less than zero, which is the objective of the optimisation.

2.2 Calculation of derivatives

The intensity at the image plane can be written as,²

$$I(X_{\text{image}}) = \frac{1}{I_F} \int_{\text{mask}} \int_{\text{mask}} dX_m dX'_m K(X_{\text{image}}, X_m) K^*(X_{\text{image}}, X'_m) T(X_m) T^*(X'_m) J(X_m, X'_m) \quad (3)$$

where I_F is the normalisation factor, $K(X_{\text{image}}, X_m)$ is the amplitude transmission factor from image to mask, $T(X_m)$ is the transmission function of the mask and $J(X_m, X'_m)$ is the mutual intensity distribution of the illuminator. $J(X_m, X'_m)$ is computed by integrating the amplitude transmission function of the condenser lens system over the illuminator aperture.

$$J(X_m, X'_m) = \int_{\text{illum.aperture}} F(X_m, X'_m, \zeta, \eta) d\zeta d\eta \quad (4)$$

where F is the amplitude transmission function of the condenser lens system, ζ and η are the coordinate variables of the illuminator aperture area.

In fact the first partial derivatives of the intensity with respect to the illuminator aperture parameters is a by-product of the intensity calculation. For example, take the case of a circular illuminator. To calculate the intensity, we have to integrate over the area of the circular aperture. To calculate the derivative with respect to the partial coherence factor σ , we have to integrate just over the circular boundary. This information is available when calculating the intensity. Indeed, $K(X_{\text{image}}, X_m)$ is given by,²

$$K(X_{\text{image}}, X_m) = \iint_{s_x^2 + s_y^2 \leq NA^2} ds_x ds_y G(X_{\text{image}}, X_m, s_x, s_y, s_z, z_L, M, \lambda) \quad (5)$$

where \hat{s} is a unit vector that points in the direction from the point in the Gaussian image plane on the optic axis to a point on the exit pupil plane, NA represents the numerical aperture on the right hand side of the objective lens, z_l is the defocus value and M is the magnification. Again for the same reason as for the illuminator parameters, we see that the first partial derivative of the intensity with respect to the numerical aperture is a by-product of the intensity calculation.

If a "flood exposure" normalisation is used for the intensity, then I_F is also a function of the stepper parameters. In the case of this normalisation, $I(X_{image}) = 1$ for the case of "flood exposure" where the mask is removed. Thus, $T(X_m) = 1$ for all X_m in (3). The equation of I_F is very similar to that of I . Therefore, the evaluation of the partial derivatives of I_F with respect to the stepper parameters is similar to the evaluation of the partial derivatives of the intensity.

The derivative of the first partial derivative of the contrast with respect to the stepper parameters can be obtained from the derivative of the intensity using (2).

2.3 Convergence

We use a conjugate gradient type algorithm to direct the iterations. The Jacobian of the cost with respect to the optimization parameters is calculated by fitting the cost function to a paraboloid based on the available information of cost and its first partial derivatives. Convergence is usually determined by the magnitude of the partial derivatives. If they are small enough, the minimum is nearby. There are also constraints on the ranges of the optimisation parameters and their magnitudes relative to each other. For example, when optimizing an annular illuminator, the individual partial coherence factors should be in the range (0, 1), and $\sigma_1 < \sigma_2$. Due to these constraints, in some of the masks at the minimum, the partial derivative of the contrast with respect to some of the optimisation parameters need not be zero.

In most cases, the optimisation takes about five to six iterations to converge. For many masks, there is significant improvement in line width control and in some cases, there is improvement in line shortening and corner rounding errors, as will be demonstrated in the examples below.

3. DEFOCUS RANGE

Defocus range is critical in the manufacturing environment and must be taken into account at all times. In the case of the optimisation of stepper parameters, it is not a good idea to minimise the sum of the contrasts at zero defocus (best focus) and end range defocus. This is because, the minimum solution may involve the situation in which the aerial image is perfect and meets almost all the specifications at one defocus, whereas at the other defocus it fails badly in meeting the specifications. OPC cannot be used to correct this kind of a mask. It is better to optimize the contrasts at the two defocuses separately and then take the average value of the optimized parameters. This tends to ensure uniform failures at both defocuses, which means that OPC can do a good job in correcting the mask.

There is a significant reduction in the amount of OPC corrections needed to repair an optimised mask as compared to the unoptimised case. This will be demonstrated in the examples.

4. EXAMPLES

1. "Nikon" test mask with circular illuminator

In this example we present the optimization of the "Nikon" mask (Shiraishi et al.). The mask has a circular illuminator. The optimization parameters are the partial coherence factor of the illuminator and the numerical aperture. In figures 1(a) and 1(b) we plot the portions of a part of the mask, before and after the optimization, which have failed to meet design specifications, as reported by our optical design rule checker (ODRC), CDLOSS.⁵ CDLOSS is used to determine where the mask pattern fails to print, assuming a simple threshold model for the exposure, post-exposure bake, and dissolution. It is clear from these figures that massive line width failures have been corrected by the optimization. However, there is no improvement in the line shortening or corner rounding errors. These failures have been calculated at 0.8μ defocus. In figure 1(c) we plot the contrast as a function of the iteration number. It is clear that the contrast decreases as the optimization proceeds and convergence is attained after about six iterations. In figure 1(d) we plot the optimization parameters NA and σ as a function of iteration number.

2. SRAM gate mask with double annular illuminator

In figure 2(a) we plot the CAD layout of an SRAM gate design. The double annular illuminator consists of two sets of annular apertures, one around the other. Thus, there are four partial coherence factors. In figures 2(b) and 2(c) we plot a portion of the failures of the SRAM mask at 0.8μ defocus, before and after optimization. It is seen that there are few remaining failures of the optimized mask. In this case there is a significant improvement in line width, line shortening and corner rounding errors. In figure 2(d) we give a table of the partial coherence factors and the corresponding contrasts as the iteration proceeds. After the second iteration σ_3 and σ_4 remain constant because of the constraint $\sigma_3 < \sigma_4$.

3. Influence of the optimization of stepper parameters on OPC corrections

In figure 3(a) we plot the CAD layout of an OPC corrected test mask with an unoptimized circular illuminator. In figure 3(b) we plot the CAD layout of an OPC corrected test mask with an optimized circular illuminator. We obtained the optimized partial coherence factor by taking the average of the optimized partial coherence factors at 0.0 and 0.4μ defocus. It can be seen that very few OPC corrections have to be made for the optimized case as compared to the unoptimized case. This is because many of the design specifications are already met after optimizing the illuminator. The OPC corrections were determined using our program OPTIMASK.^{3,4} In figures 3(c) and 3(d) we plot the aerial images of the unoptimized test mask and OPC corrected optimized test mask at 0.0 and 0.4μ defocus overlayed onto the original mask. Improvements in line shortening, line width and corner roundings can be seen. In figures 3(e) and 3(f) we plot the design failures of the OPC corrected optimized test mask. It is seen that these are minimum. These failures were determined by CDLOSS.

4. SRAM test mask with annular illuminator

In figure 4(a) we plot the OPC corrected SRAM test mask (provided by DEC) with an unoptimized annular illuminator. In figure 4(b) we plot the OPC corrected DEC SRAM mask with the optimized annular illuminator. The reduction in the OPC corrections for the mask with the optimized illuminator can be clearly seen. The aerial images of these masks were calculated at I-line exposure.

5. Comparison of total failure lengths

In table 1, we give the total length of the portions of the mask which failed to meet design specifications, for the various masks described above. In all the cases, a reduction in the failure length for the mask with the optimized illuminators is apparent.

Mask	λ (nm)	Defocus (μ)	Illum.	Total Failure Length (μ)	
				Unoptimised Illum.	Optimized Illum.
"Nikon"	365	0.8	Circle	113.85	69.82
SRAM gate	365	0.8	Double Annulus	295.7	129.7
SRAM gate	365	0.0	Double Annulus	26.7	14.38
Test mask	365	0.4	Annulus	33.77	9.69
Test mask	365	0.0	Annulus	34.33	16.828
SRAM test	365	0.4	Annulus	27.5	23.8
SRAM test	365	0.0	Annulus	44.4	18.03
SRAM test (quarter μ)	248	0.4	Annulus	23.85	14.40
SRAM test (quarter μ)	248	0.0	Annulus	23.00	14.32

Table 1: Total failure lengths for various masks.

Mask	Wavelength (nm)	Defocus (μ)	Illuminator	Minimum Contrast
SRAM test	365	0.4	Circle	4.1917
SRAM test	365	0.4	Annulus	4.1509
SRAM test	365	0.4	Double Annulus	4.1095

Table 2: Minimum contrast for various optimized illuminators.

6. Comparison of various illuminators

In table 2, we give the contrast in the aerial image for the DEC SRAM mask at 0.4μ defocus for optimized circular, annular and double annular illuminators. As can be seen from the table, the double annular illuminator does the best with the lowest contrast. The annular illuminator is next best while the circular illuminator is worst at least for this case.

5. SUMMARY

In this paper, we have described an algorithm to optimize stepper parameters. We have shown examples of how the optimisation procedure works for some current mask designs. We have also shown the influence of the optimised stepper parameters on OPC corrections.

6. ACKNOWLEDGEMENTS

This work was supported in part by the Advanced Research Projects Agency and by the Air Force Office of Scientific Research. The authors are grateful to Eric Jackson for his help in providing them with the CDLOSS and OPTIMASK results and to Dr. Len Gruber of DEC for providing them with the SRAM test mask.

7. REFERENCES

- [1] D. Cole, E. Barouch, U. Hollerbach, and S. A. Orssag, "Calculation of Aerial Images for High Numerical Aperture Projection Systems", J. Vac. Sci. Technol. B10, 3037, 1992.
- [2] D. Cole, E. Barouch, U. Hollerbach, and S. A. Orssag, "Derivation and Simulation of Higher Numerical Aperture Scalar Aerial Images", Jap. J. Appl. Phys. 31 (48) 4110, 1992.

- [3] E. Barouch, U. Hollerbach, and R. Vallishayee, "OPTIMASK: An OPC Algorithm for chrome and Phase-Shift Mask Design", *Proceedings of SPIE*, 1995.
- [4] C. A. Spence, J. I. Nistler, R. Vallishayee, E. Barouch, and U. Hollerbach, "Optimization of PSM Designs Including Defocus Effects", *Proceedings of SPIE*, 1995.
- [5] C. A. Spence, J. I. Nistler, E. Barouch, U. Hollerbach, and S. A. Orszag, "Automated Determination of CAD Layout Failures Through Focus: Experiment and Simulation", *Proceedings of SPIE* 2197 302, 1994.

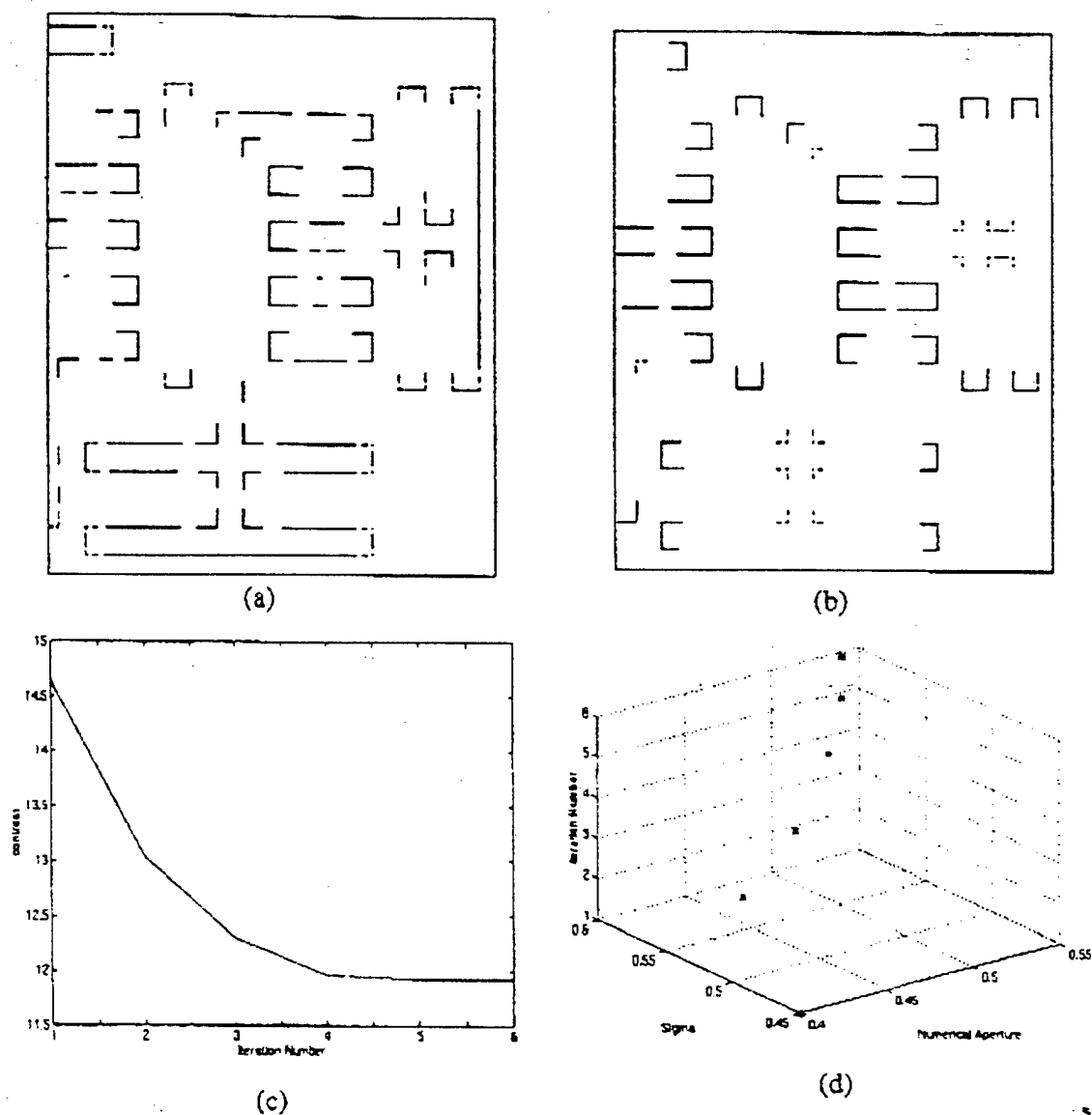
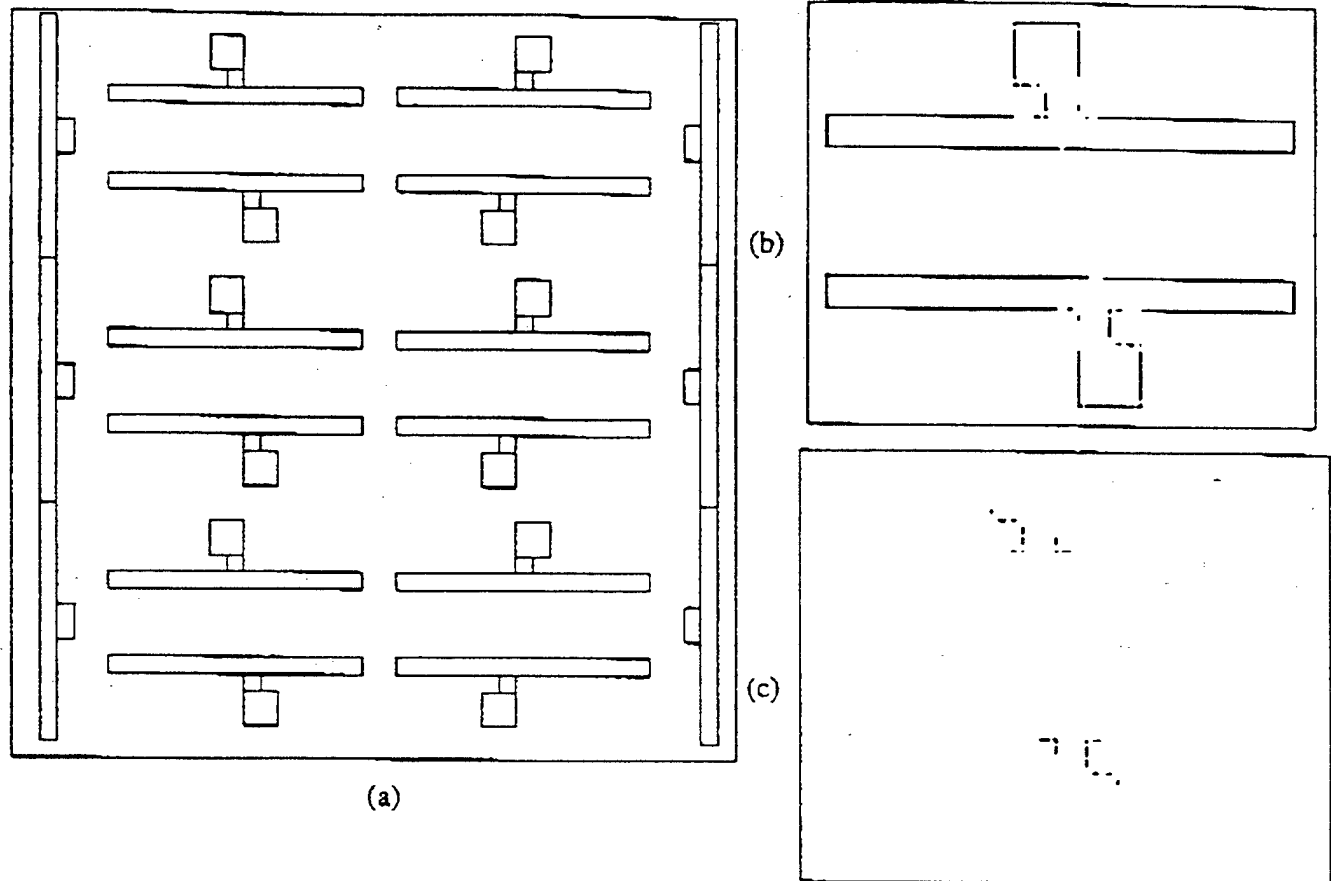


Figure 1

- (a) Portions of a part of the unoptimized "Nikon" mask that have failed to meet design specifications ($\sigma = 0.45$, $NA = 0.4$, defocus = 0.8μ)
- (b) Portions of a part of the optimized "Nikon" mask that have failed to meet design specifications ($\sigma = 0.6$, $NA = 0.54$, defocus = 0.8μ)
- (c) Contrast (cost function to be minimized) as a function of the iteration number, as the optimization proceeds
- (d) Optimization parameters (σ and NA) as a function of the iteration number as the optimization proceeds



	σ_1	σ_2	σ_3	σ_4	Contrast
1	0.4	0.5	0.6	0.7	15.479
2	0.38	0.55	0.62	0.69	15.156
3	0.36	0.60	0.61	0.63	14.876
4	0.38	0.57	0.61	0.63	14.824
5	0.39	0.56	0.61	0.63	14.811
6	0.34	0.49	0.61	0.63	14.898
7	0.41	0.51	0.61	0.63	14.789

(d)

Figure 2

- (a) CAD drawing of SRAM mask
- (b) Portions of a part of the unoptimized SRAM mask that have failed to meet design specifications ($\sigma_1 = 0.4$, $\sigma_2 = 0.5$, $\sigma_3 = 0.6$, $\sigma_4 = 0.7$, $NA = 0.54$, defocus = 0.8μ)
- (c) Portions of a part of the optimized SRAM mask that have failed to meet design specifications ($\sigma_1 = 0.41$, $\sigma_2 = 0.51$, $\sigma_3 = 0.61$, $\sigma_4 = 0.63$, $NA = 0.54$, defocus = 0.8μ)
- (d) Table showing how the contrast of the SRAM mask changes with the optimization parameters ($\sigma_1, \sigma_2, \sigma_3, \sigma_4$) as the optimization proceeds ($NA = 0.54$, defocus = 0.8μ)

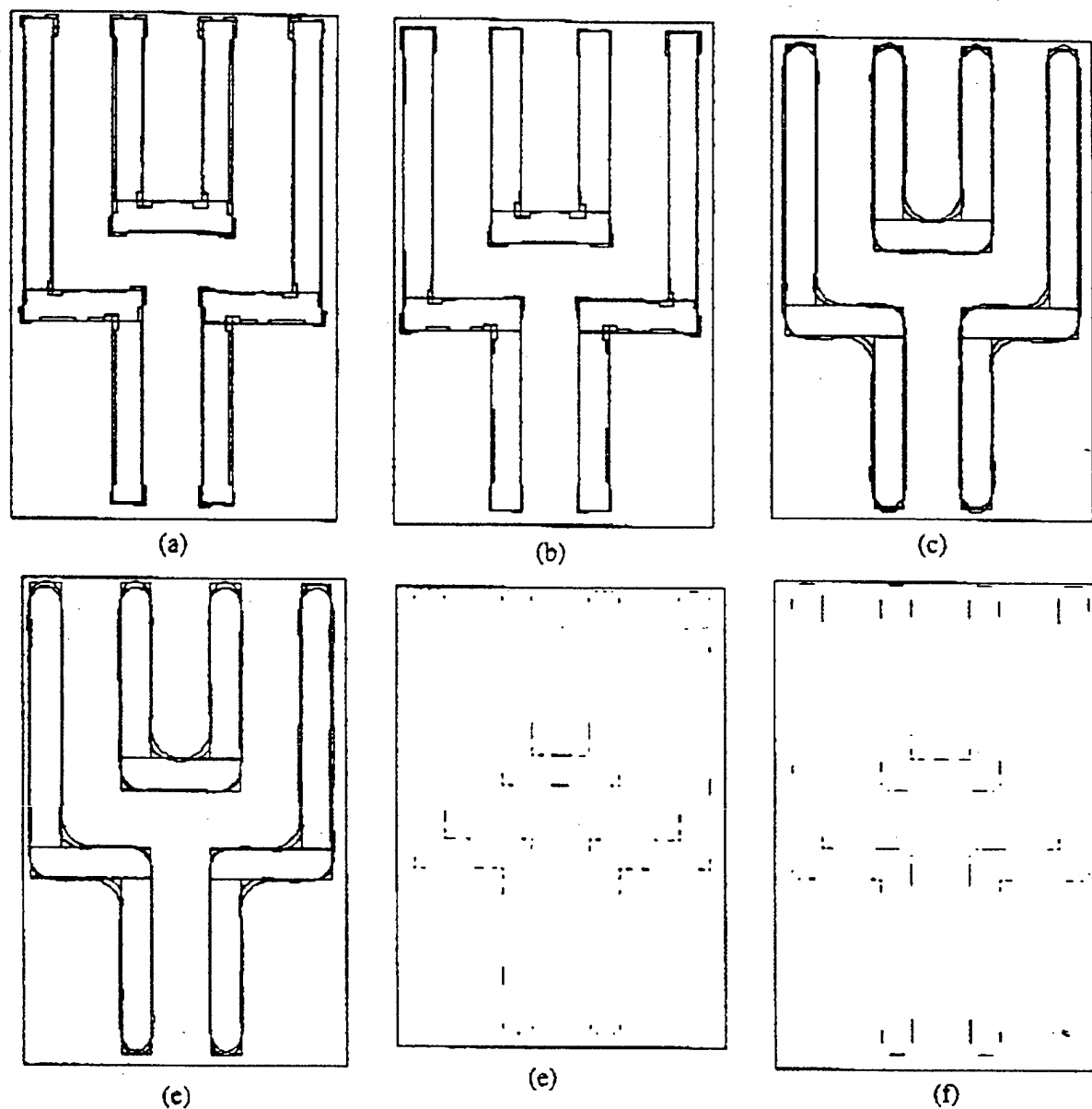
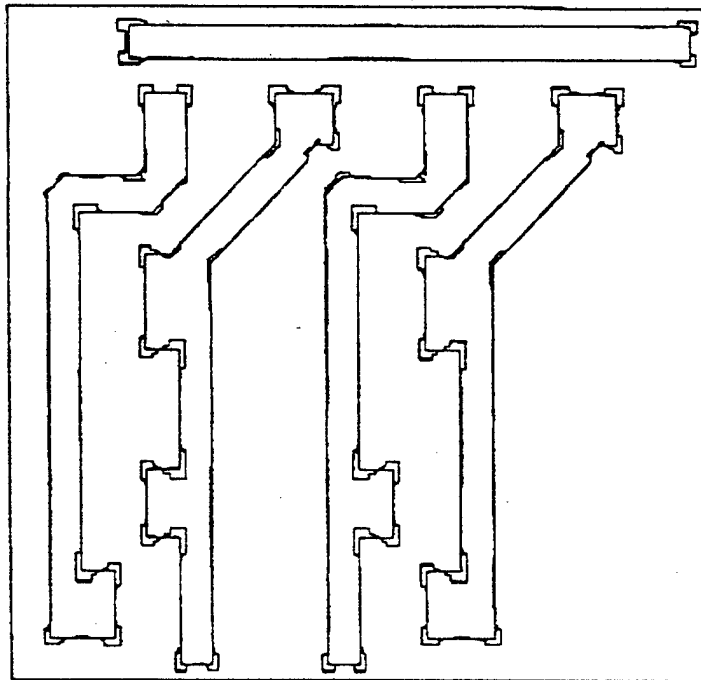
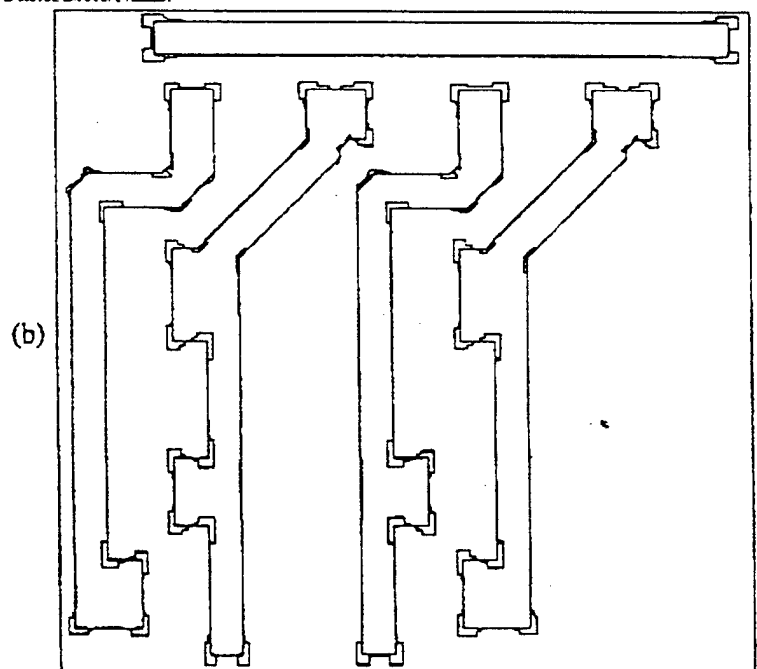


Figure 3

- (a) OPC corrected mask with unoptimized circular illuminator ($\sigma = 0.48$, $NA = 0.54$)
- (b) OPC corrected mask with optimized circular illuminator ($\sigma = 0.62$, $NA = 0.54$)
- (c) 0.3 intensity level contour of aerial images of unoptimized mask and OPC corrected optimized mask overlayed onto original mask (defocus = 0.4μ)
- (d) 0.3 intensity level contour of aerial images of unoptimized mask and OPC corrected optimized mask overlayed onto original mask (defocus = 0.0μ)
- (e) Portions of original mask which are out of specs. at 0.4μ defocus for the optimized OPC corrected mask
- (f) Portions of original mask which are out of specs. at 0.0μ defocus for the optimized OPC corrected mask



(a)



(b)

Figure 4

(a) OPC corrected DEC SRAM mask with unoptimized annular illuminator.

(b) OPC corrected DEC SRAM mask with optimized annular illuminator.

Advanced Transformational Analysis Applied to E-Beam Proximity Effect Correction

E. Barouch⁺, R. Coifman^{*}, J. Ma^{*},
M.C. Peckerar⁺⁺, and V. Rokhlin^{*}

⁺ *Department of Applied and Computational Mathematics
Princeton University
Princeton, NJ 08544*

^{*} *Yale University
Department of Mathematical Sciences
New Haven, CT 06473*

⁺⁺ *U.S. Naval Research Laboratory
Nanoelectronics Processing Facility
Washington, DC 20375*

February 1993

Abstract

In this paper we address the problem of dose correction in the data bases consistent with ultra-large-scale integration. It is shown that recent advances in transformation theory provide a natural platform on which to build these dose correctors. Specifically, transformation approaches making use of compactly supported, smooth basis functions are shown to be particularly suitable. This is a natural result of the evolution of mathematically based correctors currently in use. Previous work of Parikh, MacDonald and others employed global transform method to determine the values of "corrected" dose.

In most cases, the mathematical inversion is essentially ill posed, in other words, the exact pattern desired cannot be obtained using a finite Gaussian sum.

In this paper a set of smooth basis elements of compact support are employed. The mathematically smooth form of the basis makes it easy to match doses at boundaries without Gibbs phenomenon. Thus the transform field can be partitioned for optimum speed. Consequently, while most transformation complexities are of order N^6 (the inversion of an $N^2 \times N^2$ matrix) where N^2 is the number of grid points characterizing the database, we developed an algorithm of complexity $N^2 \log N$. A method of dose field bias is employed to stem the requests for negative dose.

The heart of the numerical process is essentially based on an adapted fast non-uniform-grid Fourier Transform combined with proper filtering and geometric localization methods. Several examples are given.

1 Introduction

In E-beam lithography the exposure tool emits a very narrow electron beam (as low as 5nm spot-size) at each prescribed pixel. It deposits its energy in the resist. However, due to forward scattering, backward scattering as well as tertiary reactions, the very narrow energy distribution changes its nature. This fact maybe characterized by expressing the deposited "point spread (PS) energy" (the energy deposited in the resist per pixel of the E-beam writer), as a narrow forward peak f , with added corrections. The main correction that must be taken into account is the backward scattering b . For simplicity of presentation we limit ourselves to back scattering correction only, even though in principle any addition to the forward peak can be included.

In essence we are adopting the classic model of Kyser and Murata [1] to represent the PS function as a sum of two Gaussians. The formulation and analysis below is independent of this fact, but it is as convenient as well as conventional representation. Following Gerber [2], Parkovich [3], Parikh [4] and Haslam and McDonald [5] we define the "Proximity Equation" as the equation to find the "Beam Density" D , which represents the machine setting to yield the desired deposited exposure energy in the resist E .

However, as correctly stated by Gerber [2] and others, the proximity equation is ill-posed. Accordingly a precise mathematical solution that makes physical sense is not possible. Heuristically it is clear. An ideal deposited energy for mask fabrication is composed of features, represented as discontinuous step function distributions. The point spread functions are infinitely differentiable. Thus any finite linear combination would yield a continuous function, rendering the problem insolvable.

The ill posedness of the proximity equation is vividly displayed when the Gaussian model is imposed [2] and is a resultant of the ill posedness of the backward diffusion problem exhibited in the forward peak.

It is the purpose of this paper to introduce a somewhat modified proximity equation and provide numerical analysis of its structure as well as some analytical insights and specific example.

Once the dose is agreed upon, we demonstrate, as examples, the profiles for simulated problems using E-beam into chemically amplified resist. Hollerbach has modified our simulator MCP [6] so resist images can be obtained for E-beam lithography, and we used the modified MCP throughout this paper. It should be mentioned that Moniwa, Yamaguchi and Okazaki [7] have introduced a 3D simulator SEED 3D.

2 Formulation and Technical Issues

We model the point spread function P as

$$P = f + \eta b. \quad (1)$$

Let b , the back scattering energy distribution, be represented as a convolution of a density B with f , i.e.

$$b = f * B \quad (2)$$

where $*$ represents a convolution, η the relative strength of the back scattering and P is obtained as

$$P = f * (1 + \eta B). \quad (3)$$

Let D represent the "energy density", or the E-beam writer settings. Thus, the total deposited energy E on the resist surface is given by

$$E = P * D = f * D * (1 + \eta B). \quad (4)$$

Let M represent the "mask function" with

$$M = \begin{cases} 0 & \text{at chromed regions} \\ 1 & \text{at transparent regions.} \end{cases} \quad (5)$$

The illposedness of the problem assures us that the proximity equation $M = E$ has no solution.

In reality, even if there is no back scattering, one CANNOT achieve better energy density (projected aerial image) than the forward-peak, and probably somewhat worse. We therefore "filter" the mask function. Specifically we assert that the best one can obtain is the mask function convolved with some filter h , somewhat wider than the forward peak f .

We now introduce the modified proximity equation:

$$M * h = f * D * (1 + \eta B) = M * f * d \quad (6)$$

with d representing the excess filter width. Taking Fourier transform of (6) yields

$$\bar{D} = \frac{\bar{M}\bar{h}}{\bar{f} + \eta\bar{b}} = \frac{\bar{M}\bar{d}}{1 + \eta\bar{B}}. \quad (7)$$

The division by $1 + \eta\bar{B}$ and the performance of the inverse transformation is another manifestation of the ill posedness of the problem.

To illustrate this point assume $\bar{d} = 1$ and expand the denominator $1 + \eta\bar{B}$ as a sum of geometric series, yielding

$$\bar{D} \simeq \bar{M}(1 - \eta\bar{B} + \eta^2\bar{B}^2 \pm \dots). \quad (8)$$

Taking the inverse transform yields the density D as the mask function minus the small but positive definite convolution $M * B$.

$$D \cong M - \eta M * B - \eta^2 M * B * B = \dots \quad (9)$$

However, M as a mask function contains domains in which it vanishes. Since $M * B$ is positive definite, D obtains negative values which are of course nonphysical. This non-physical result is a consequence of the assumption that even the modified proximity equation admits a "good" solution. The inverse transform of (7) suffers from the same fate exhibited at (9). It should be noted that the above analysis is general and not dependent on the Gaussian or any other model. The formal solution of the modified proximity equation is given as

$$D = \text{Inverse}[\bar{M}\bar{d}(1 + \eta\bar{B})^{-1}] \quad (10)$$

with the understanding that D may become negative.

The question is how can the solution D be employed? Two obvious answers are:

(1) "Shift". That means add a constant dose to D in each pixel, such that it is definite positive. This method is potentially more successful for positive resists than negative chemically amplified ones.

(2) "Chop". That means set D to vanish everywhere it obtains a negative solution.

In both cases D must be convolved with the PS function to achieve the desired energy distribution deposited in the resist.

Since D refers to machine setting it is natural to question which of these two options is better and under what general circumstances. Furthermore, what is the "best" way to shift a little and chop the rest?

It is clear that these questions must be answered coupled to the resist processing, are resist dependent and vary from system to system. In this paper we demonstrate the solution of the modified proximity equation for simple features as well as the prototype "NIKON MASK"[8]. Both the shifted and chopped versions of the resultant energy distributions are considered. Last but not least, the "mixed" algorithm is under active consideration.

3 Examples

We begin our discussion by specifying the explicit form of the PS function P . Following Kyser [1] we model the unnormalized point spread function P as at x_i, y_j by

$$P = \exp(-r^2/\beta_f^2) + \eta(\beta_f/\beta_b)^2 \exp(-r^2/\beta_b^2) \quad (11)$$

with

$$r^2 = (x - x_i)^2 + (y - y_j)^2 \quad (12)$$

and β_f and β_b represent the half width of the forward and backward scattering peaks, and η as before represents the ratio of the number of backscattered to forward scattered electrons at the depth of interest.

It should be emphasized that we present β_f and β_b as constants for convenience only. In general they are functions of depth (z) and can be treated as such with no extra complication. A detailed discussion is available in reference [9].

The parameter β_f, β_b and η are taken from reference [9] and [7], and we use for the first example the notation $\sigma_j^2 = 2\beta_j^2$.

The first example demonstrates graphically the "shifted" algorithm for a $1-d$ mask with 4 features. Figure 1a displays the point spread function, fig. 1b the Gaussian filter, fig. 1c the targeted and "filtered" mask and fig. 1d the density. In this case the shift is about 4%. It is clear that the variable size test features can be better resolved the further they are from each other.

We now turn our attention to the generic "Nikon Mask"(NM) reported in ref. [8]. It is a logic-memory cell of $10.15 \mu \times 14.7 \mu$ and it contains 31 features. The feature sizes are all $.35 \mu$. The mask is illustrated in fig. 2 and its SEM in ref [8]. In this paper we extract the information on the line $y = 5$ as well as discussing the whole mask.

The parameters used in this example are displayed in Table 1.

Table 1				
Energy	β_f	β_b	η	Source
10Kev	.22	.65	.51	ref 9
25Kev	.06	2.6	.51	ref 9
30KeV	.21	3.84	.70	ref 7
40KeV	.04	6.0	.42	ref 9

When the aerial image of the NM is computed the minimum resolution is 64×96 . For the purpose of this paper we have chosen 256×256 grid for all cases involving the NM.

Figure 3 presents level curves as a color distribution of the "chopped" density at 25keV, while fig. 4a shows the same data file in a relief surface form and fig. 4b offers a sectioned display of the same surface.

Figures 5a-5d compare the beam intensity for 10keV and 25keV with both the "chopped" and "shifted" versions. As discussed earlier, the "intensity" represents the actual energy deposited in the resist during the process of E-BEAM exposure.

All the remaining examples are concentrated on the cut line $y = 5$. In fig. 6a, 6b we offer the extract of the actual solution of the modified proximity equation at 30keV as well as 40keV. The sharpness of the forward peak at 40keV is believed to be responsible for the fast oscillations near the nominal mask edge. Furthermore, fig. 7a-7d, 8a-8d corresponds to the actual intensity for both chopped and shifted cases for all energies considered in Table 1. It is clear that for a fast, high contrast negative resist, the chopped is uniformly superior. However, the situation is somewhat more vague when slow positive resist is used requiring further analysis.

We now turn our attention to the resist simulation process. The "chop-25keV" energy profile was selected for the actual exposure simulation, using SAL-601 resist layer $.5\mu$ thick. The acid production of the negative, chemically amplified resist was modelled with the classic Dill's equation, employing the modified MCP[6]. The post exposure bake(PEB) was carried out with MCP's system of reaction diffusion equations and the dissolution was simulated using

MCP's Hamilton-Jacobi algorithm.

In fig. 9a,9b we compare the results of the simulation for the line $y = 5$ for 40sec and 60sec PEB and 60sec dissolution time. At 60sec the lines have not opened up as yet but are expected to open-up for longer dissolution time. On the other hand, the 40sec PEB profiles open up very neatly, providing near optimal conditions for 0.35μ feature size E-beam lithography. In fig. 10 we exhibit a comparison of the "chopped" and "shifted" 10keV and 25keV final profile after 60sec PEB. Only the "chopped 25keV" is a viable candidate for manufacturing a $1 \times$ Nikon Mask at $.35\mu$ feature size.

4 Summary and Conclusions

We have demonstrated that 0.35μ $1 \times$ mask is feasible to manufacture with SAL-601 resist. With optimization of resist processing coupled to an optimization scheme for density selection, we expect to develop an algorithm providing 0.25μ feature size masks.

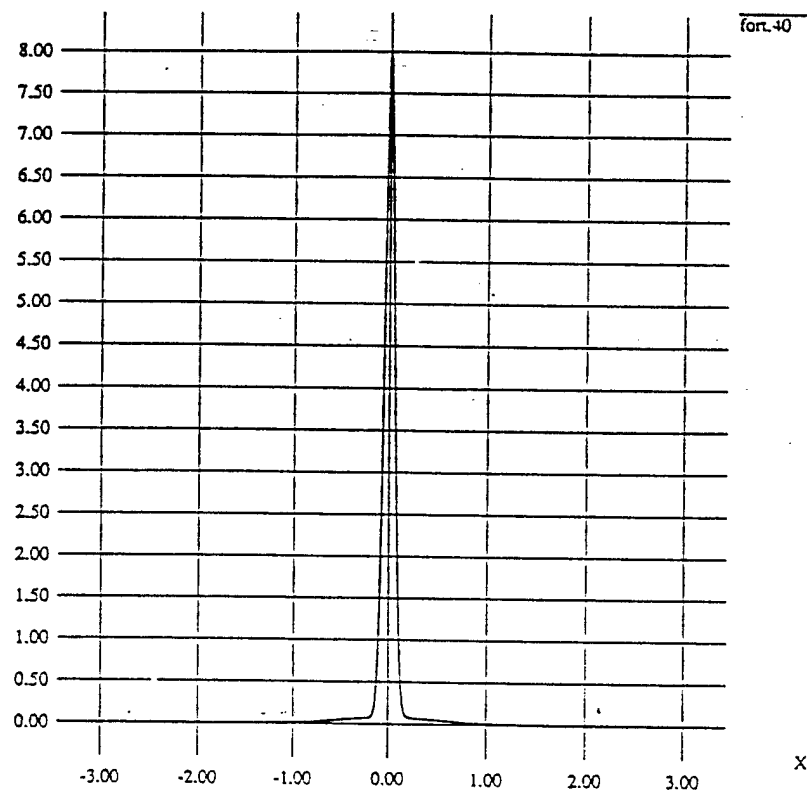
5 Acknowledgements

The authors express their gratitude to U. Hollerbach for his assistance in performing the resist simulations. This study has been supported by AFOSR(EB) and DARPA(EB JM and MCP).

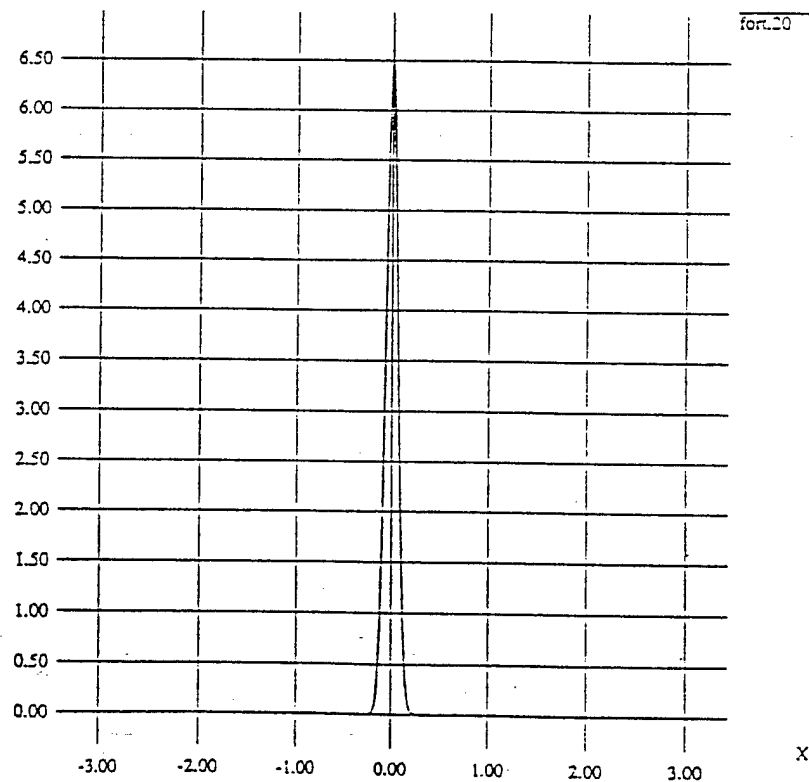
References

- [1] D. Kyser and K. Murata, Proc. 6th Int. Conf. Electron Ion Beam Science and Technology, 1974.
- [2] P.D. Gerber, J. Vac. Sci. Tech. B6, 432 (1988).
- [3] J.M. Pavkovich, J. Vac. Sci. Tech. B4, 159 (1986).
- [4] M. Parikh, IBM J. Res. Dev., 24, 438 (1980).
- [5] M.E. Haslam and J.F. McDonald, J. Vac. Sci. Tech., B4, 168 (1986).
- [6] U. Hollerbach, Ph.D. Thesis, Princeton University, 1992. See also E. Barouch, J.W. Cahn, U. Hollerbach and S.A. Orszag, Jour. of Sci. Comp., 6, 229 (1991).
- [7] A. Moniwa, H. Yamaguchi and S. Okazaki, EIPB Issue, J. Vac. Sci. Tech., B10 (1992).
- [8] N. Shiraishi, S. Hirukawa, Y. Takeuchi and N. Magome, Proc. SPIE 1674, 741 (1992).
- [9] S.P. Murarka and M.C. Peckerar, "Electronic Materials", Academic Press Pub. Co., San Diego (1989).

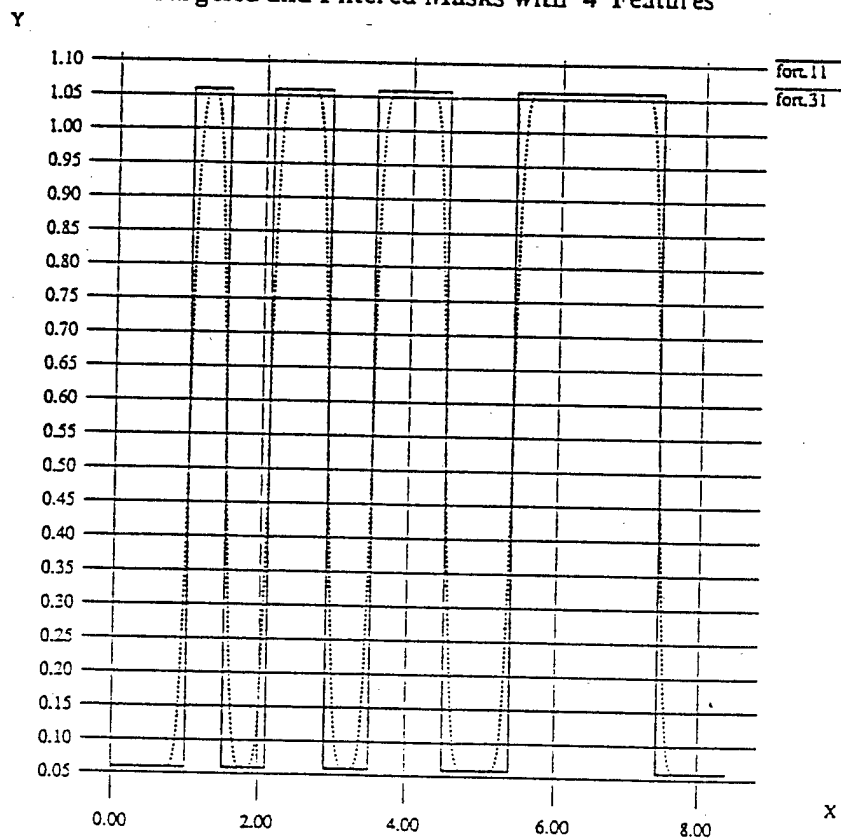
Spread with Two Gaussian Bells (SIGMA = 0.05, 0.5) (ETA = 1, .1)



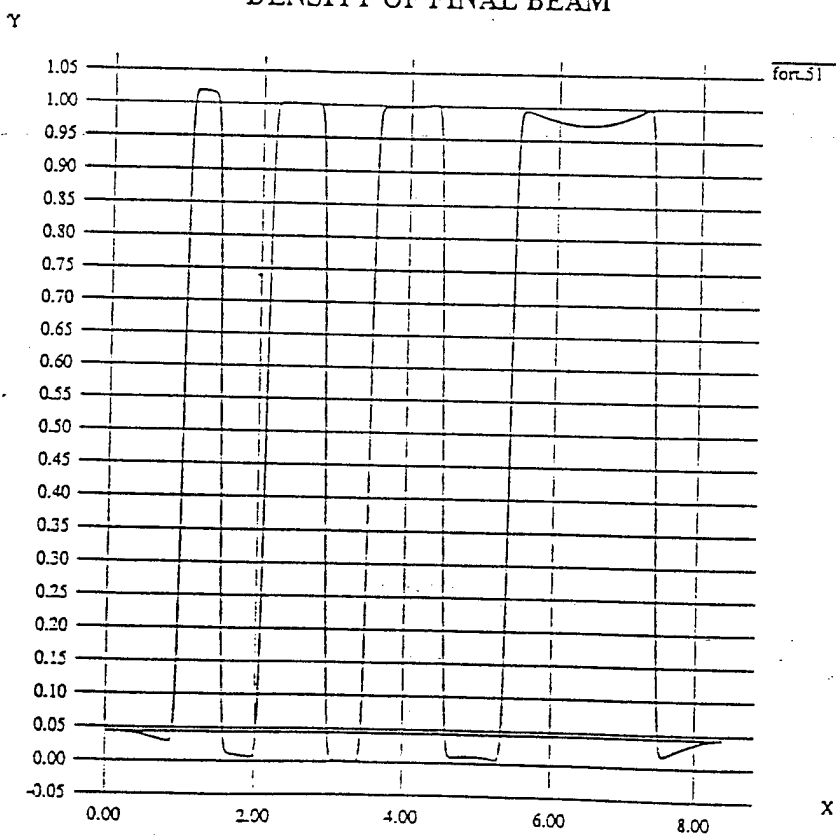
GAUSSIAN FILTER (SIGMA = 0.06)



Targeted and Filtered Masks with 4 Features

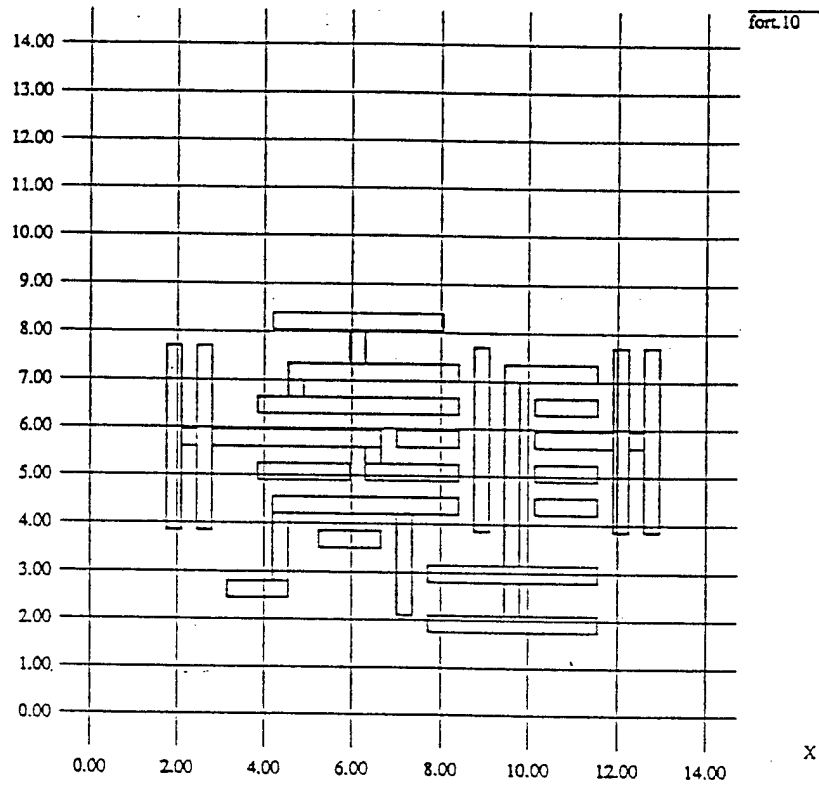


DENSITY OF FINAL BEAM

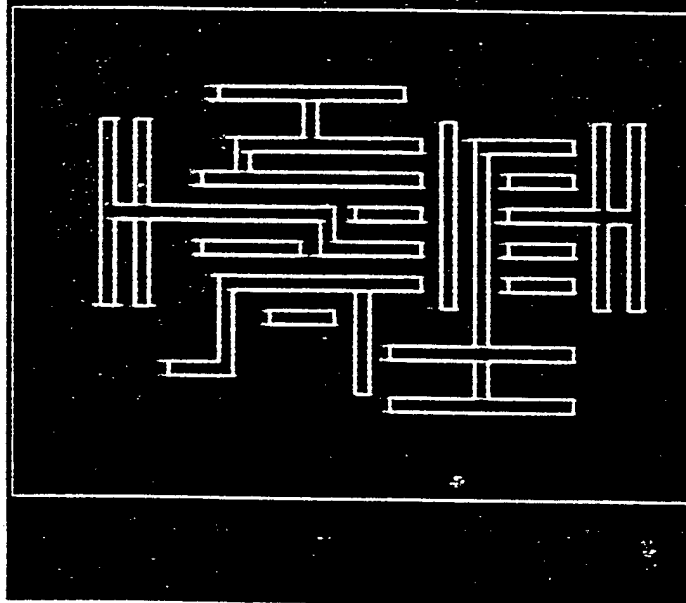


Nikon Mask (# of features= 31)

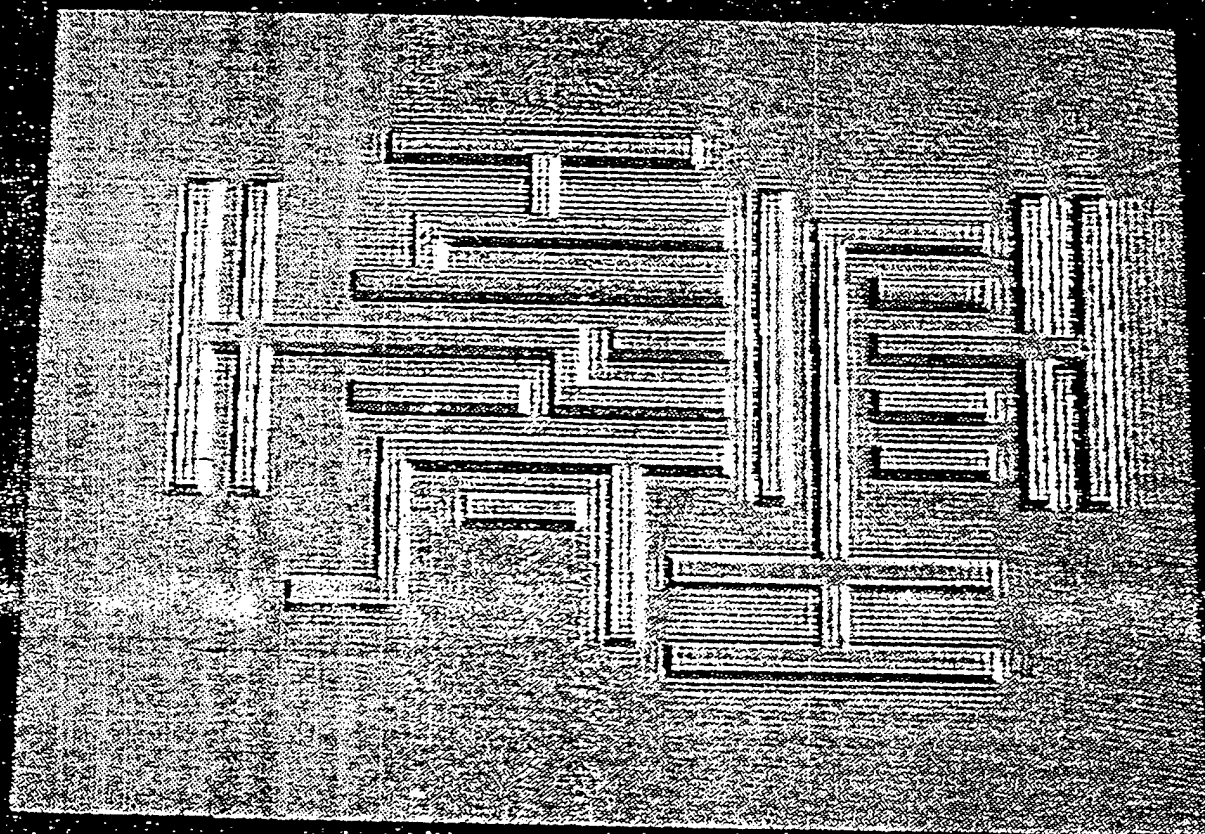
Y

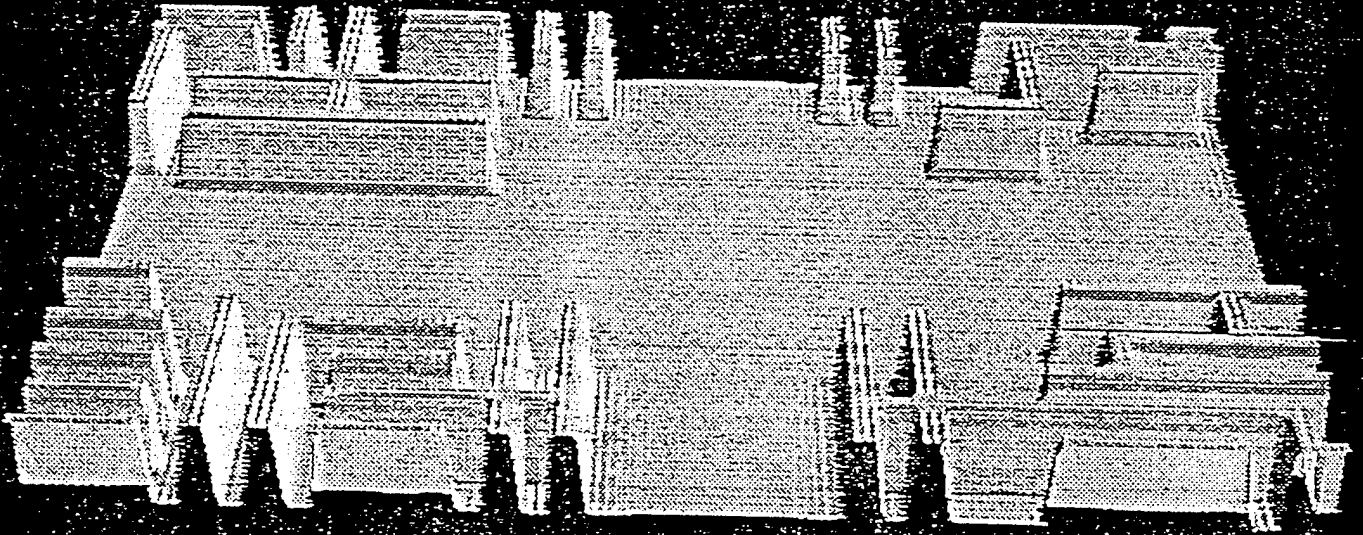


Nikon Mask by J.H.Ma

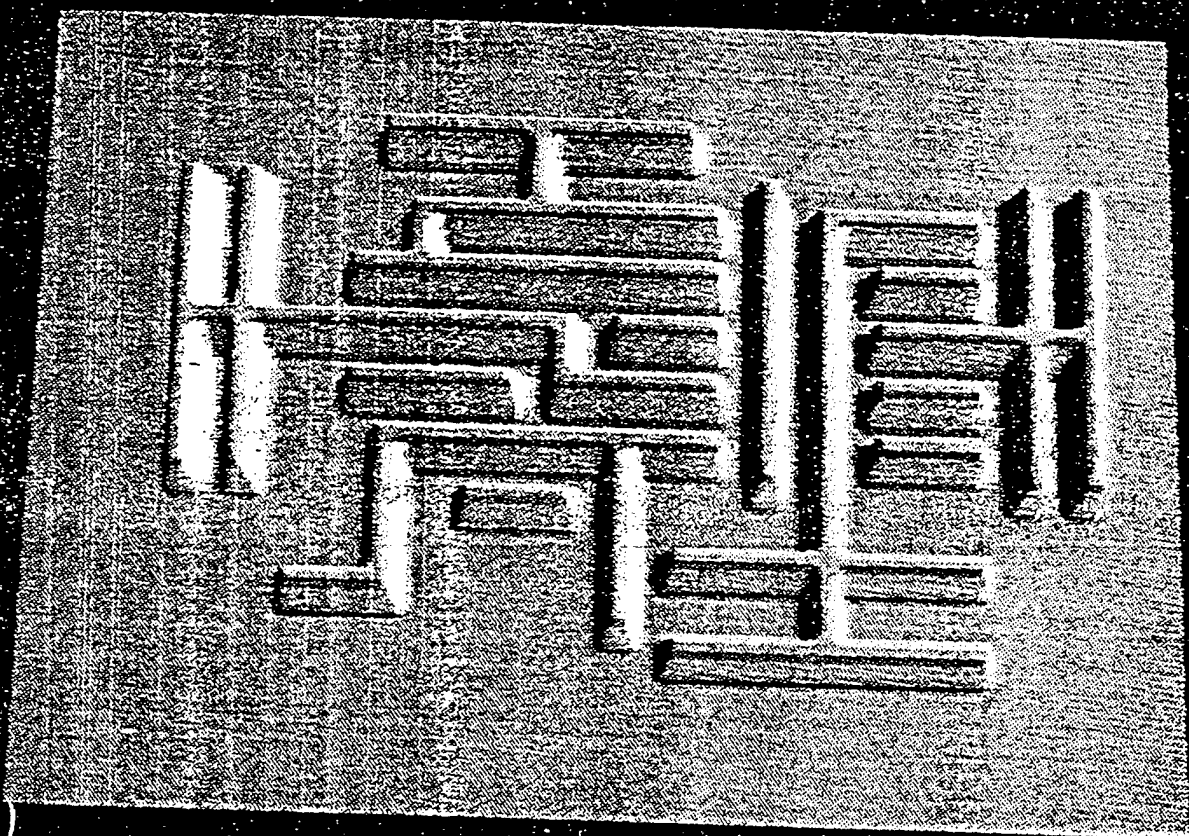


beam-density-shifted 25kev

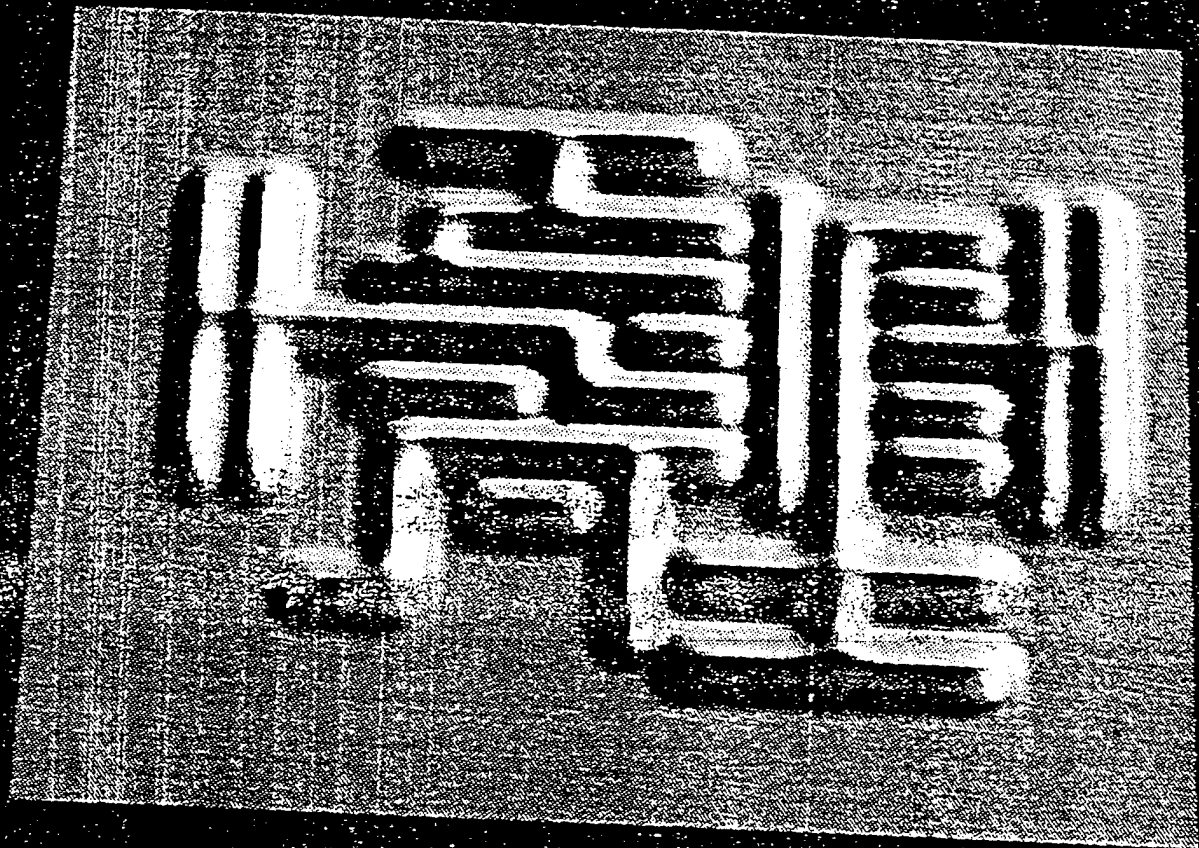




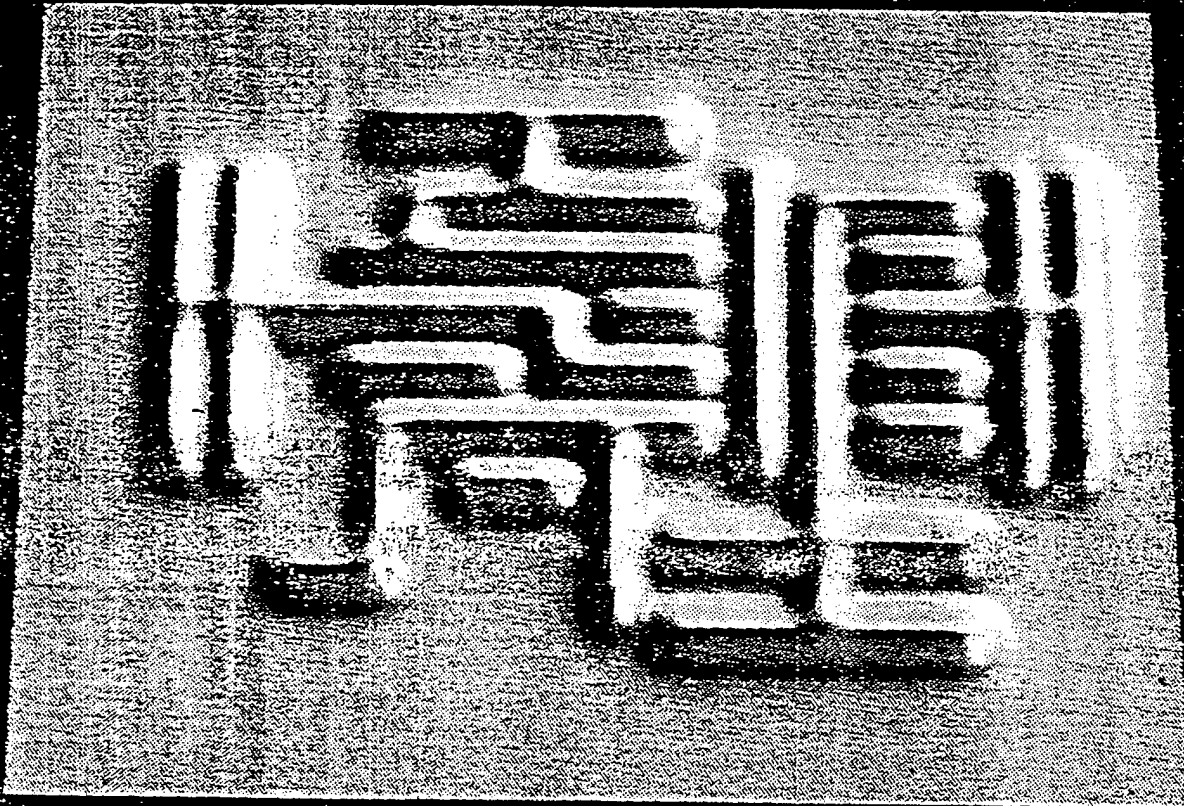
beam-intensity-shifted 25kev



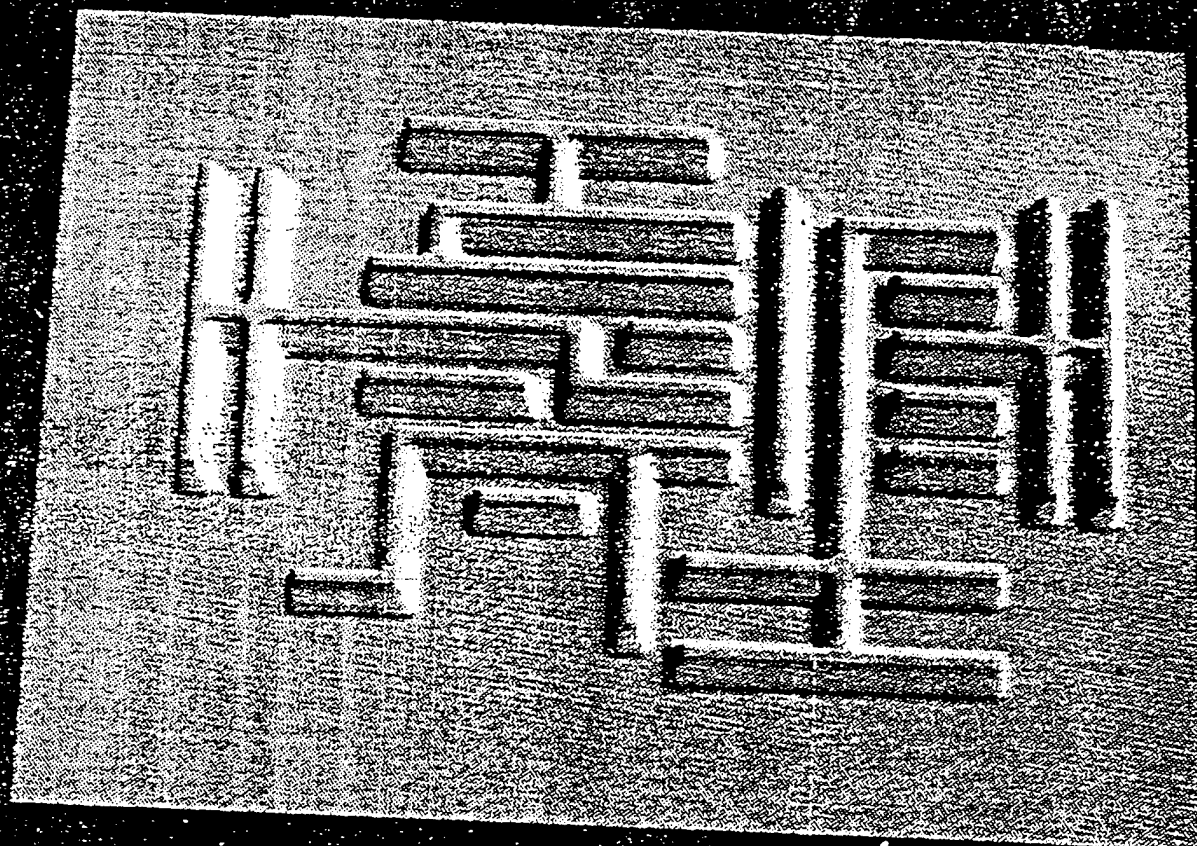
beam-intensity-shifted 10kev



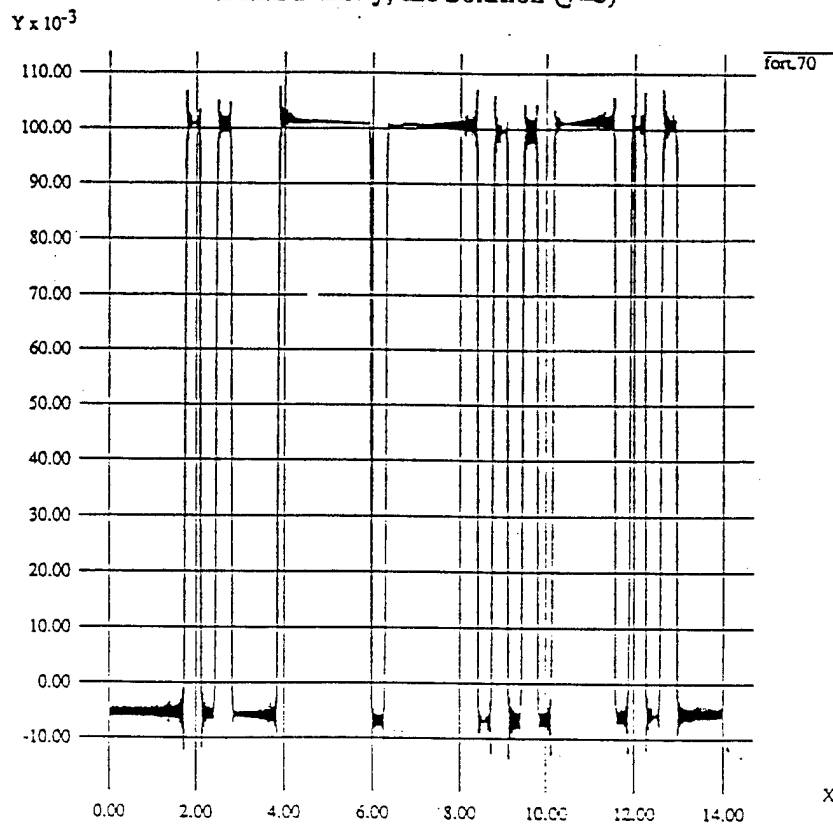
beam-intensity-chopped-10kev



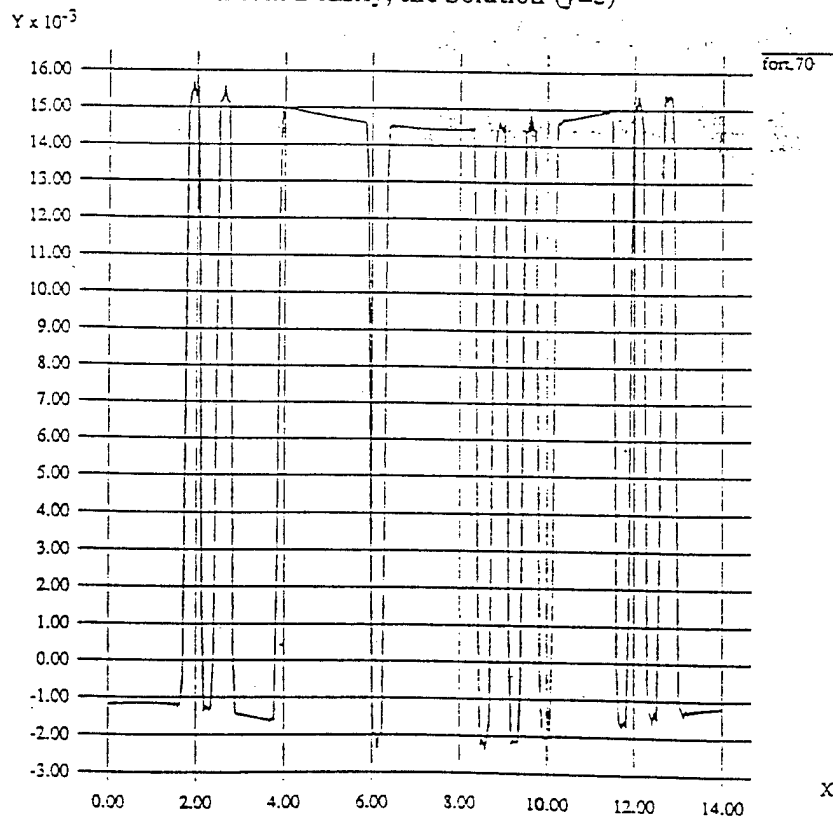
beam-intensity-chopped.25kev



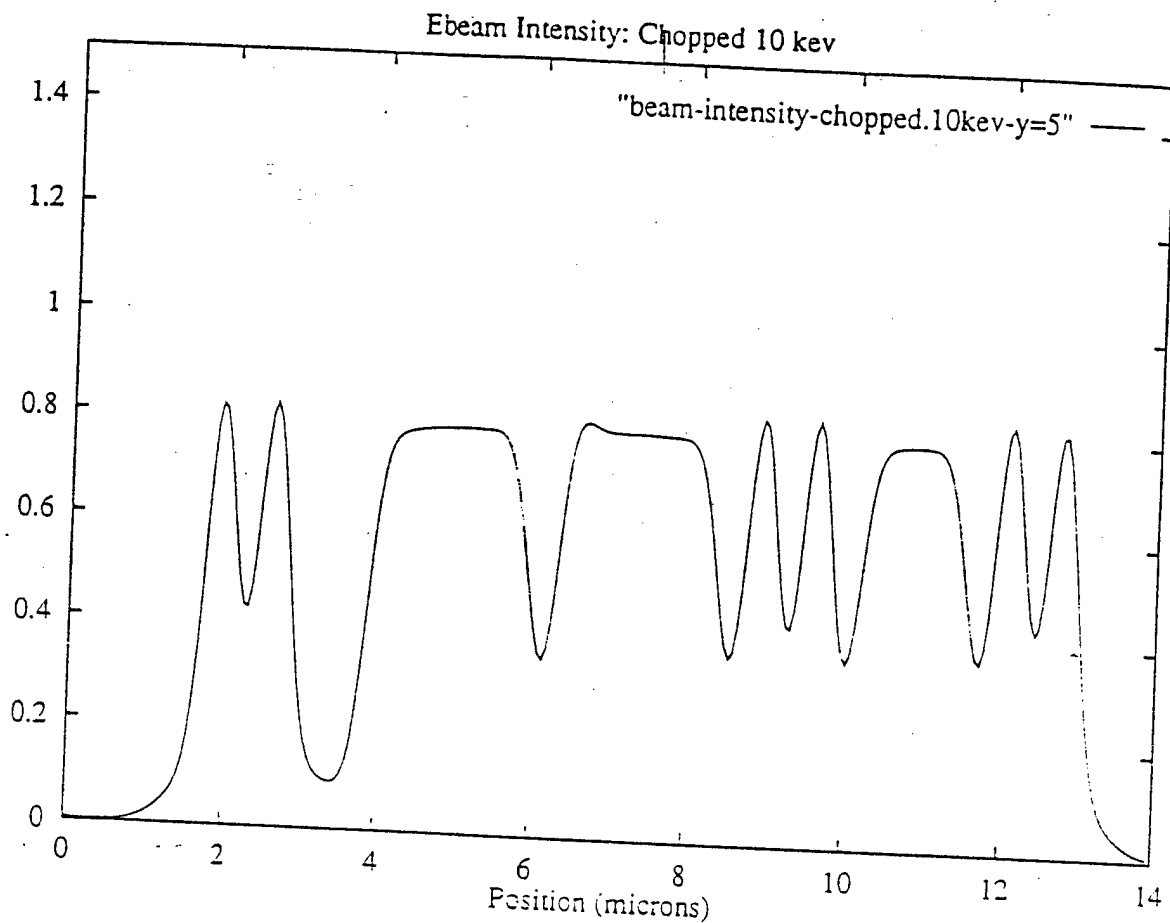
Beam Density, the Solution ($\gamma=5$)



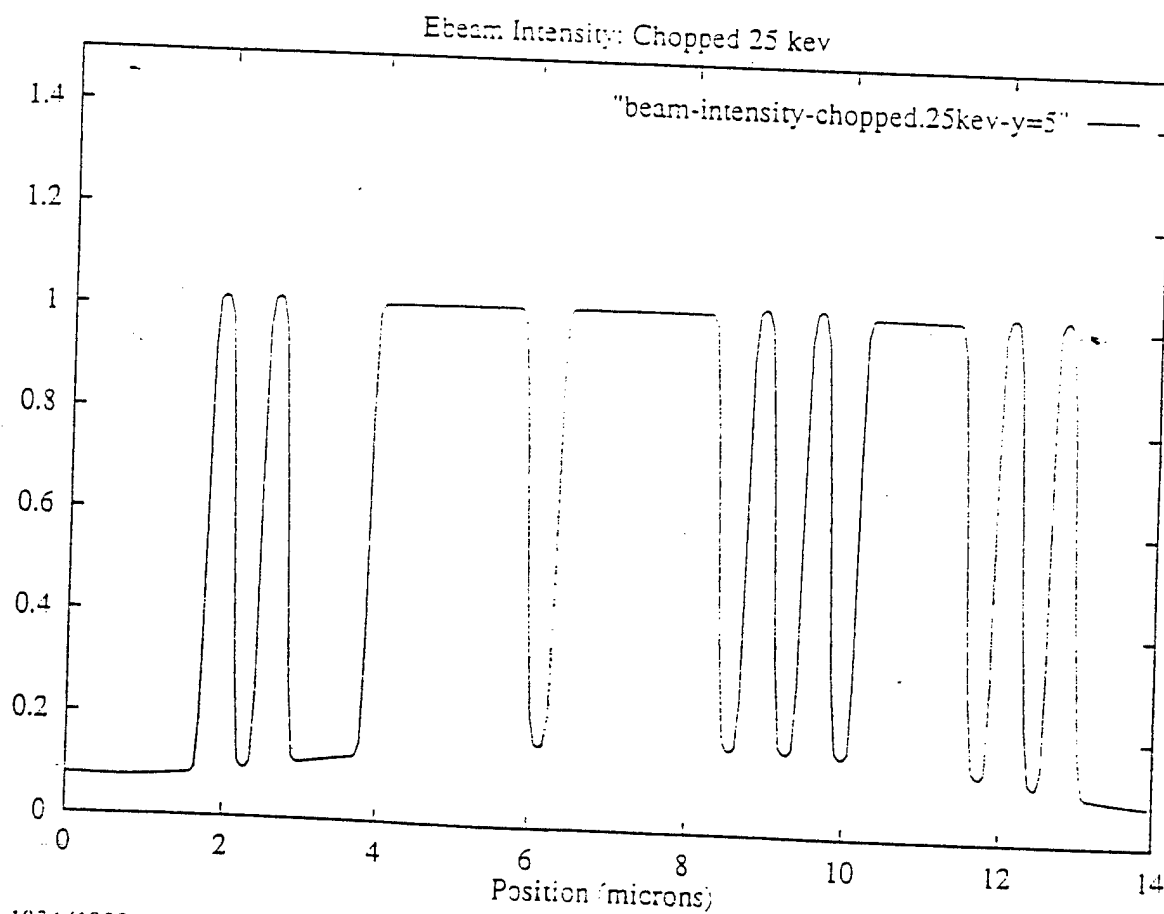
Beam Density, the Solution ($\gamma=5$)

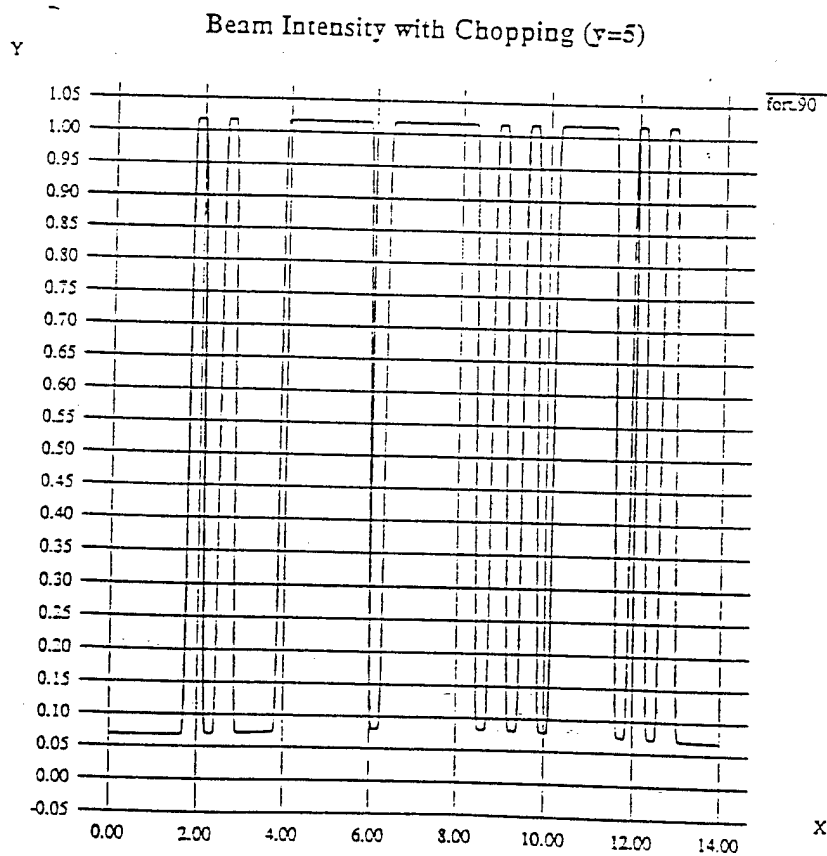
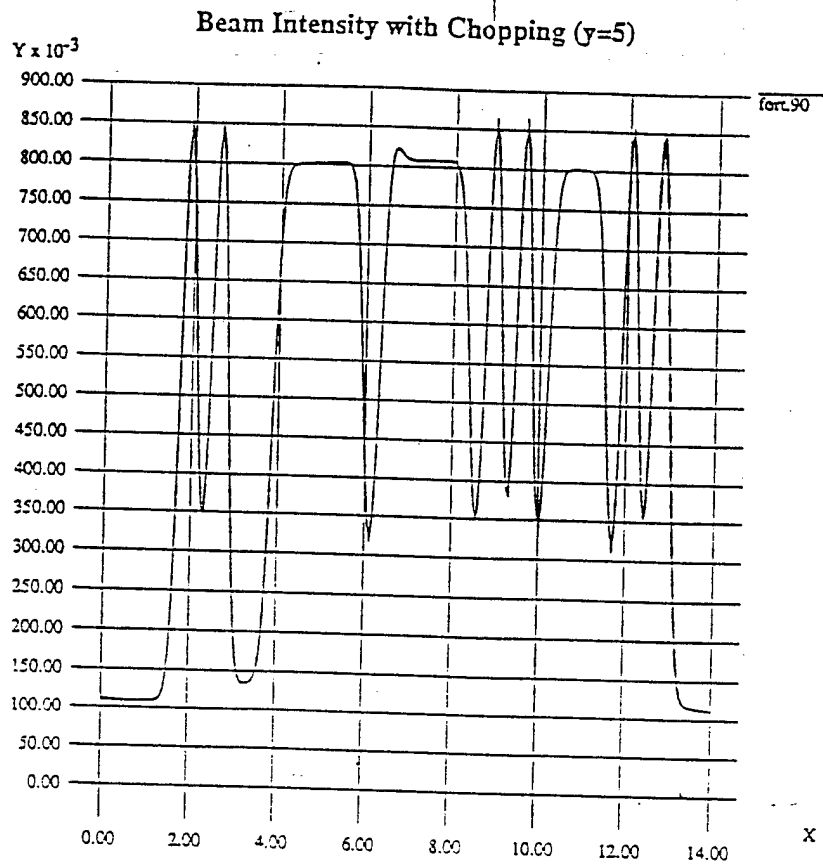


Relative Intensity

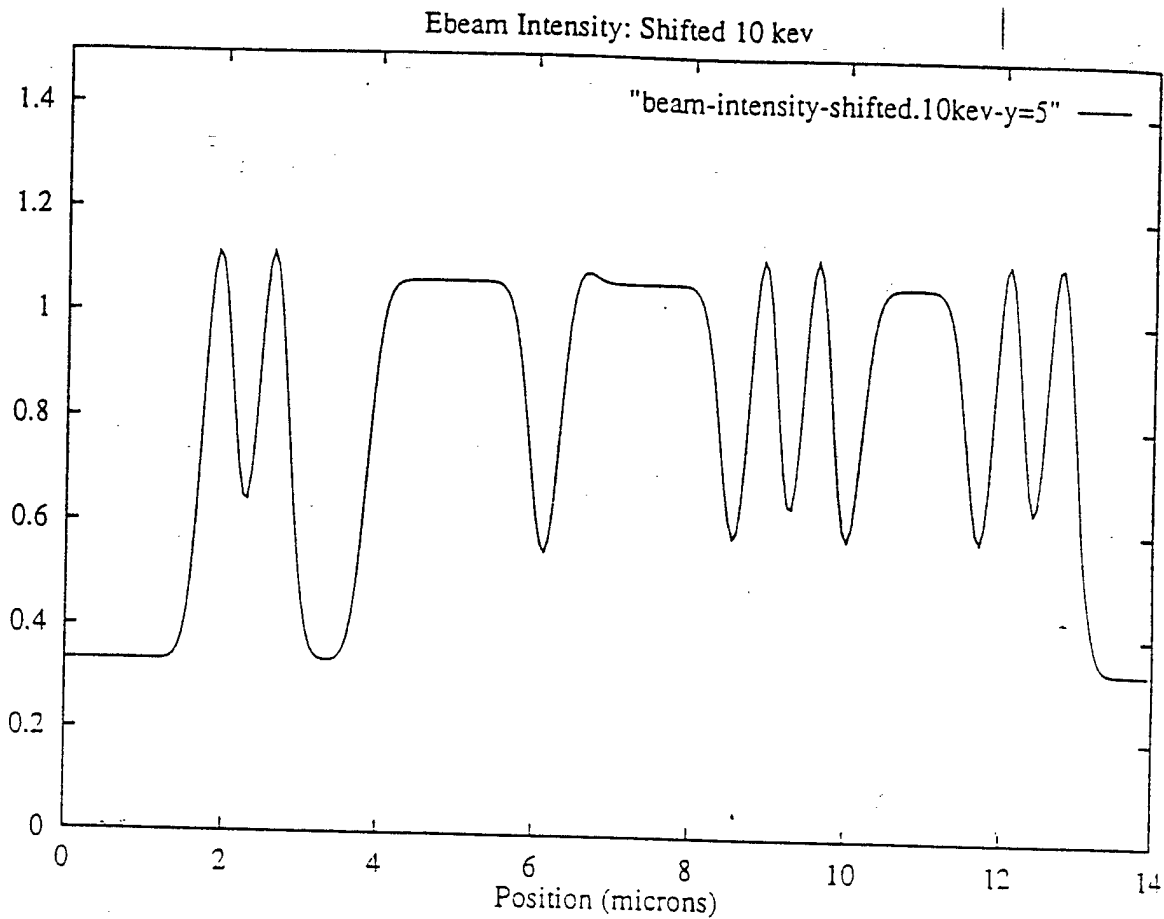


Relative Intensity

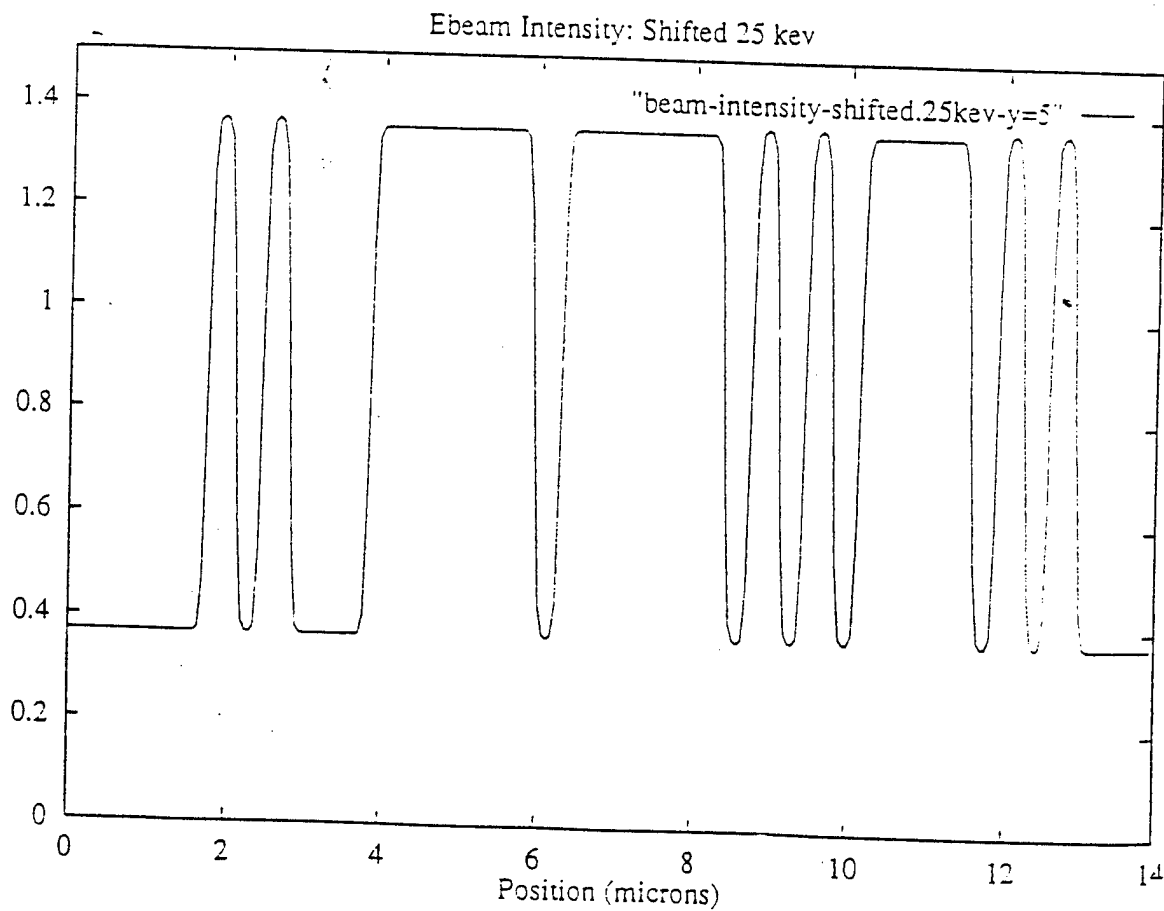




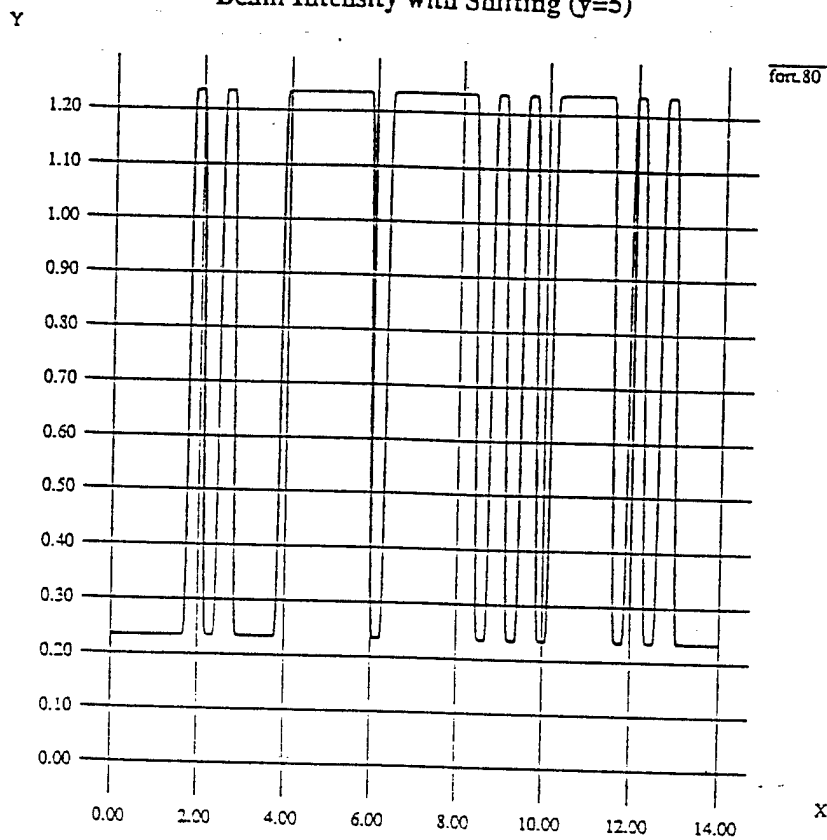
Relative Intensity



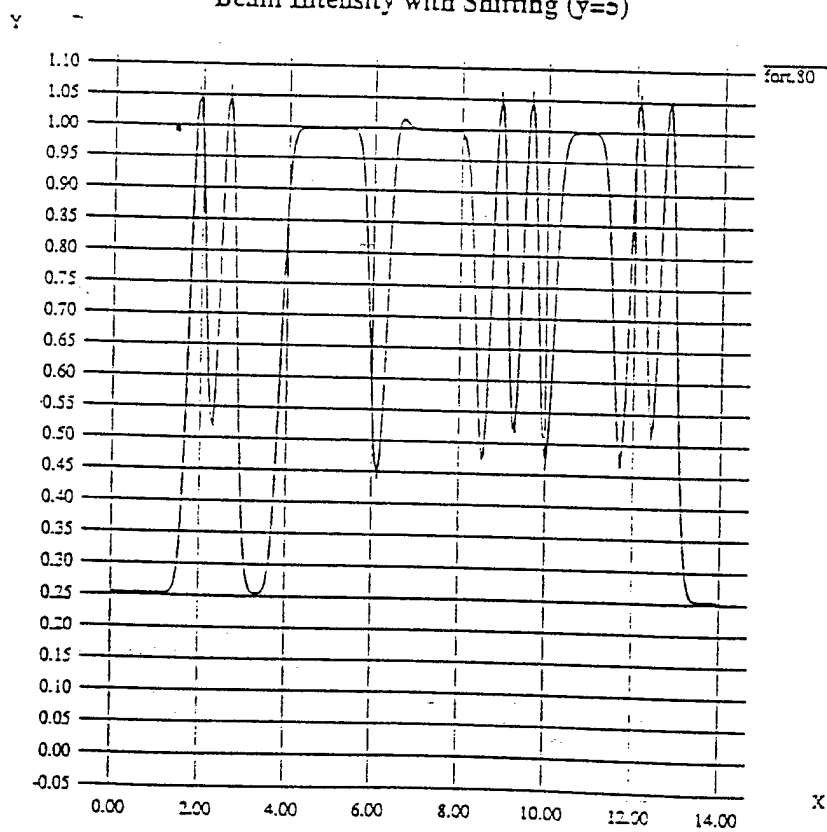
Relative Intensity



Beam Intensity with Shifting ($y=5$)



Beam Intensity with Shifting ($y=5$)



Chop - 25 kev

Acid Concentration: 0-3%



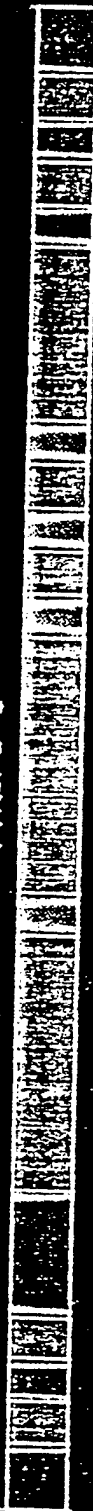
Cross-Link Concentration After 40 sec PBB: 0-1



Final Dissolved Profile After 40 sec PBB



Cross-Link Concentration After 60 sec PBB: 0-1



Final Dissolved Profile After 60 sec PBB

